

Open Research Online

The Open University's repository of research publications and other research outputs

The rate of profit as a random variable

Thesis

How to cite:

Wells, Patrick Julian (2007). The rate of profit as a random variable. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2007 Patrick Julian Wells

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000574b>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

The rate of profit as a random variable

Thesis submitted in accordance with the requirements of
The Open University
for the degree of Doctor of Philosophy

by

Patrick Julian Wells
BSc (Econ) MSc

School of Management
The Open University

July 2007

Abstract

The rate of profit as a random variable

This thesis is a systematic attempt to investigate two conjectures about the distribution of company rates of profit: that it should be log-normal (Gibrat 1931), and that it should be gamma distributed (Farjoun and Machover 1983).

A large set of company accounts data is analysed, and partial support found for Gibrat and for a generalised version of Farjoun and Machover.

The analysis includes a demonstration of different empirical distributions for different profit rate measures, a demonstration of power law tails in all measures of the profit rate, and a demonstration of size effects (differences in tail weights) in financial ratios. Annual variation in the overall skewness and kurtosis of profit rate distributions is shown to be dominated by variation in the power law tails.

L-moments, a recent innovation in robust methods to deal with extreme values, are used in conjunction with a size-weighted sampling scheme to identify possible models for distributions of the profit rate at the capital level.

Farjoun and Machover derive their hypothesis from a particular conception of the process of capitalist competition. A rival conception, that of Glick (1985), is tested using company accounts data and shown to be vulnerable to criticism concerning the scope of its data set, the test statistic employed, and its choice of profit rate measure. More fundamentally, it is also dependent on doubtful premises about the within-industry distribution of profit rates, as *L*-moment analysis demonstrates.

Acknowledgements

I am indebted to my supervisors, Janette Rutterford and Andrew Trigg, for their supportive and constructive criticism of the earlier drafts of chapters; in addition, Andrew kindly scrutinized the bulk of the thesis.

I acknowledge the School of Management of The Open University for its support of my research.

I am indebted to many fellow scholars who have provided a stimulating intellectual environment in which to conduct this study; in particular I must mention helpful discussions with Alan Freeman, Andrew Kliman, Moshé Machover and Ian Wright.

Participants in the mini-conferences of the International Working Group on Value Theory, and in the conferences of the Association for Heterodox Economics and of the European Association for Evolutionary Economics, also provided encouraging responses to a number of presentations relating to parts of the work.

Nothing would have been possible without the very practical input of Roberto Simonetti, who made available data derived from the Financial Analysis Made Easy company accounts database. A number of members of the S-Plus and R discussion lists provided help at crucial moments, among whom John Hosking, Chris Jones and Brian Ripley are noteworthy.

Finally, I must record my gratitude to the late Nina Youngman, who inspired me to set out on the road that led to this project, and to Rose Clements, without whom I would not have completed it.

Table of contents

List of Tables and Figures vi

Chapter 1 Introduction I

- 1.1 The importance of the rate of profit 2
 - 1.1.1 Which rate of profit? 4
- 1.2 The rate of profit as a random variable 6
 - 1.2.1 The transformation problem and the rate of profit as a random variable 9
- 1.3 Objectives 10
- 1.4 Plan of work 12

Chapter 2 Measuring the rate of profit 14

- 2.1 Can profit rates be calculated from accounting data? 15
 - 2.1.1 Empirical relevance of IRR versus ARR 16
 - 2.1.2 Concepts of income and capital 18
 - 2.1.3 Capitalists and accounting 20
 - 2.1.4 Summary 22
- 2.2 Issues in the rate of profit 22
 - 2.2.1 Stock or flow measures? 28
 - 2.2.2 Unproductive labour and capital 29
 - 2.2.3 Interest and taxes (broad versus narrow measures) 32
 - 2.2.4 Non-capitalist sector 41
 - 2.2.5 Money or hours 42
 - 2.2.6 Housing 43
 - 2.2.7 Are workers' taxes variable capital? 43
 - 2.2.8 Current versus historic cost 44
- 2.3 Conclusion 47
 - 2.3.1 Use of accounting data 47
 - 2.3.2 Profit rate concepts 47

Chapter 3 Data and exploratory analysis 50

- 3.1 Data source and construction of profit rate measures 50
 - 3.1.1 Data source: Financial Analysis Made Easy (FAME) 50
 - 3.1.2 Construction of profit rate measures from FAME 52
- 3.2 Empirical densities and summary statistics 59
 - 3.2.1 Estimating empirical density functions 59
 - 3.2.2 Histogram estimates of empirical profit rate densities 61
 - 3.2.3 Summary statistics 68
- 3.3 Extreme values and firm size 73
- 3.4 Conclusion 81

Chapter 4 Methods 83

4.1 Extreme values, outliers, and discordancy 84

- 4.1.1 Review of issues 84
- 4.1.2 Sources of outliers in profit rate data 87
- 4.1.3 Strategies for outliers in profit rate data 91

4.2 Robustness: the method of *L*-moments 96

- 4.2.1 Definition and description 96
- 4.2.2 *L*-moments and order statistics 96
- 4.2.3 *L*-moments and traditional moments 98
- 4.2.4 *L*-moment ratio diagrams 100
- 4.2.5 Parameter estimation using *L*-moment ratios 101
- 4.2.6 Model selection using *L*-moment ratios 102
- 4.2.7 *L*-moment summary 102
- 4.2.8 Application of *L*-moments to profit rate data at the firm level 103

4.3 Estimating capital-level profit rate distributions 104

- 4.3.1 Randomly-sized random samples 104
- 4.3.2 Empirical densities and summary statistics for capital-level data 106

4.4 Model selection with capital-level data 116

4.5 Conclusion 117

Chapter 5 The distribution of the rate of profit 119

5.1 Farjoun and Machover's probabilistic political economy 121

- 5.1.1 Statistical mechanics, the transformation problem, and the gamma distribution 121
- 5.1.2 Farjoun and Machover on Marx 127
- 5.1.3 Marx's probabilistic political economy 129
- 5.1.4 Testing Farjoun and Machover 134

5.2 Distributions of rates of return 138

- 5.2.1 *L*-moment ratio analysis of profit rate distributions 138
- 5.2.2 Sampling properties of RS2 estimation 145

5.3 Zipf plot analysis of tails of profit rate distributions 152

- 5.3.1 Power-law tails and rates of profit 153
- 5.3.2 Comparison of power-law tails between profit rate measures 157
- 5.3.3 Annual variation in power-law tails 165
- 5.3.4 Summary of Zipf plot analysis 175

5.4 Conclusion 176

Chapter 6 Profit rates and the competitive process 178

6.1 Glick on gravitation of profit rates 180

6.2 Critique of Glick 185

- 6.2.1 Analysis is at too high a level of aggregation 187

6.2.2	The preferred statistic for measuring gravitation is questionable	189
6.2.3	The full implications of the hypothesis are not tested	191
6.2.4	No attention is paid to the <i>shape</i> of profit rate distributions	192
6.3	Reproducing Glick	193
6.3.1	Methodology	193
6.3.2	Results	195
6.4	Testing Glick	198
6.4.1	The level of aggregation	198
6.4.2	The measure of dispersion	200
6.4.3	Alternative profit rate measures	201
6.4.4	Conclusion	204
6.5	Industry distributions: the L-moments perspective	204
6.5.1	Methods	205
6.5.2	Results	208
6.5.3	Resolving power of L-moments	218
6.6	Conclusion	223
Chapter 7	Conclusions	224
7.1	Contributions	225
7.1.1	Evaluating profit rate distributions	225
7.1.2	Longitudinal variation in profit rate distributions	227
7.1.3	Competition as gravitation	228
7.1.4	L-moments in economics	230
7.1.5	Choice of profit rate definition	232
7.3	Directions for future research	235
Appendix	237	
Probability density function	237	
Two-parameter (scale and shape)	237	
Three-parameter (location, scale and shape)	238	
Four-parameter (location, scale, shape and power/shape)	239	
The generalised gamma and other distributions	239	
Other properties	241	
Characterisation	241	
Inequality measures	241	
Reproductive property	241	
Estimation	242	
Bibliography	244	

List of Tables and Figures

Table 2.1: conceptual issues in Moseley (1991).....	25
Table 2.2: profit rate measures to be tested	49
Table 3.1: definitions of FAME variables.....	53
Table 3.2: construction of profit rate measures using FAME variables	56
Figure 3.1: income and asset quadrant.....	57
Table 3.3: firms with negative net assets.....	58
Figure 3.2: empirical density estimates for Gillman1; histogram and kernel methods.....	60
Figure3.3: empirical density functions, Gillman’s ‘marxian’ measures.....	62
Figure3.4: empirical density functions, Gillman’s ‘capitalist’ measures.....	63
Figure3.5: empirical density functions, accounting ratios	64
Figure 3.6 (a): empirical density functions, Glick measures.....	65
Figure 3.6 (b): empirical density functions, Glick measures	66
Table 3.4 (a): summary statistics for underlying data, Gillman profit rate measures.....	69
Table 3.4 (b): summary statistics for underlying data, accounting ratios.....	70
Table 3.4 (c): summary statistics for underlying data, Glick profit rate measures.....	71
Table 3.5: tail quantiles of firm-level distributions	74
Figure 3.7: log ranges of profit rate measures; extreme bars truncated.....	74
Figure 3.8: firm size and range of profit rate, Gillman’s ‘marxian’ measures	76
Figure 3.9: firm size and range of profit rate, Gillman’s ‘capitalist’ measures	77
Figure 3.10: firm size and range of profit rate, accounting ratios	78
Figure 3.11: firm size and range of profit rate, Glick measures	79
Figure 4.1: extremes, outliers and contaminants (Barnett and Lewis, 1994: 9)	85
Figure 4.2: firm size distributions, selected capital measures (see text)	89
Figure 4.3: loci of selected distributions in L -skewness, L -kurtosis space (γ_3 , γ_4 space)	100
Figure 4.4: γ_3 , γ_4 space chart for 21 profit rate measures; see Table 4.1 for key.....	103
Table 4.1: key to Figure 4.4.....	104
Figure 4.5: empirical density estimate for Gillman 1, metasample of weighted data	107

Figure 4.6: empirical density estimate for Gillman's 'marxian' measures, metasample of weighted data	109
Figure 4.7: empirical density estimate for Gillman's 'capitalist' measures, metasample of weighted data	110
Figure 4.8: empirical density estimate for accounting ratios, metasample of weighted data	111
Figure 4.9: empirical density estimate for Glick measures, metasample of weighted data ..	112
Table 4.2 (a): summary statistics for firm-weighted profit rate distributions, Gillman's 'marxian' measures.....	114
Table 4.2 (b): summary statistics for firm-weighted profit rate distributions, Gillman's 'capitalist measures.....	114
Table 4.2 (c): summary statistics for firm-weighted profit rate distributions, accounting ratios.....	114
Table 4.2 (d): summary statistics for firm-weighted profit rate distributions, Glick measures	115
Figure 4.10: log ranges of capital-weighted profit rate distributions.....	115
Figure 4.11: ${}_3, {}_4$ space chart for weighted profit rate measures; see Table 4.1 for key	116
Table 5.1: testing Farjoun and Machover (1983).....	137
Figure5.1: L -skewness and L -kurtosis of Gillman marxian measures	139
Figure5.2: L -skewness and L -kurtosis of Gillman capitalist measures:	140
Figure5.3: L -skewness and L -kurtosis of accounting ratios	141
Figure5.4: L -skewness and L -kurtosis of Glick measures.....	143
Figure5.5: L -skewness and L -kurtosis of Gillman 4 in closer focus	145
Figure5.6: Gillman 4, 1995; L -skewness and L -kurtosis of RS2 samples.....	146
Figure5.7: Gillman 4, 1995; L -moment space of RS2 samples	149
Figure5.8: Zipf plot of samples from exponential samples compared to the Pareto distribution	153
Figure5.9: Zipf plot of unweighted and weighted Gillman 4 measures, with gamma distribution	154
Figure5.10: Zipf plot of annual variation in weighted Gillman 4 measures.....	155
Figure5.11(a): Zipf plot analysis, unweighted Gillman profit rate measures, marxian and capitalist.....	158
Figure5.11(b): Zipf plot analysis, unweighted accounting ratios.....	159

Figure 5.11(a): Zipf plot analysis, unweighted Glick profit rate measures	160
Figure 5.12(a): Zipf plot analysis, size-weighted Gillman profit rate measures, marxian and capitalist.....	162
Figure 5.12(b): Zipf plot analysis, size-weighted accounting ratios	163
Figure 5.12(c): Zipf plot analysis, size-weighted Glick profit rate measures	164
Figure 5.13(a): Zipf plot analysis of annual variation; Gillman marxian measures.....	166
Figure 5.13(b): Zipf plot analysis of annual variation; Gillman marxian measures.....	167
Figure 5.14(a): Zipf plot analysis of annual variation; accounting ratios	169
Figure 5.14(b): Zipf plot analysis of annual variation; accounting ratios	170
Figure 5.15(a): Zipf plot analysis of annual variation; Glick measures	172
Figure 5.15(b): Zipf plot analysis of annual variation; Glick measures.....	173
Table 5.2: tail quantiles (excerpted from Table 3.5)	175
Table 6.1: Glick's Value Line statistics.....	184
Table 6.2: rank order of Glick's Value Line results by V_1	185
Table 6.3: Glick Value Line results ranked by relative gravitation	191
Figure 6.1: histogram of industry average rates of profit, after Figure 4.4 in Glick (1985)	192
Table 6.4: mapping from UK SIC codes to Glick manufacturing industries	196
Figure 6.2: \mathbf{v}_i statistics for two-digit industries, Glick profit rate measures	197
Figure 6.3: \mathbf{v}_i statistics for different levels of aggregation, Glick profit rate measures	199
Figure 6.4: $\hat{\mathbf{v}}$ statistics for two-digit industries, Glick profit rate measures	200
Figure 6.5: $\hat{\mathbf{v}}$ statistics for different levels of aggregation, Glick profit rate measures	200
Figure 6.6: \mathbf{v}_i statistics for different levels of aggregation, multiple profit rate measures....	202
Figure 6.7: $\hat{\mathbf{v}}$ statistics for different levels of aggregation, multiple profit rate measures....	203
Figure 6.8: Gillman 3, mean/ cLv space	210
Figure 6.9: Gillman 4, mean/ cLv space	210
Figure 6.10: ATO, mean/ cLv space.....	211
Figure 6.11: Glick 1, mean/ cLv space.....	211
Figure 6.12: Glick 5, mean/ cLv space.....	212

Figure 6.13: Gillman 3, L -skewness/ L -kurtosis space.....	214
Figure 6.14: Gillman 4, L -skewness/ L -kurtosis space.....	214
Figure 6.15: asset turnover (ATO), L -skewness/ L -kurtosis space.....	215
Figure 6.16: Glick 1, L -skewness/ L -kurtosis space	215
Figure 6.17: Glick 5, L -skewness/ L -kurtosis space	216
Figure 6.18: Glick 1, L -skewness/ cLv space.....	217
Figure 6.19: Glick 5, L -skewness/ cLv space.....	217
Figure 6.20: asset turnover rate (ATO), whole economy, firm-level distribution.....	218
Figure 6.21: ATO firm-level distribution by one-digit sector.....	219
Figure 6.22: ATO firm-level distribution; sector 8 by two-digit industry.....	220
Figure 6.23: ATO, L -skewness/ L -kurtosis space, sector 8	221
Figure 6.24: ATO, L -skewness/ cLv space, industries 81-85	222
Figure 6.25: ATO, mean/ cLv space, industries 81-85	222
Figure A.1: gamma distributions; shape parameters as indicated in legend	238
Figure A.2: relationship of generalised gamma to other distributions (from McDonald 1984)	240

Chapter I Introduction

This work is a contribution to the empirical investigation of the Marxist profit rate and its supposed equalisation. We use a very large set of U.K. company accounts data to identify and estimate probability density functions describing the empirical distribution of profit rates. Key theoretical works on this topic are those of Gibrat (1931) and Farjoun and Machover (1983), and we present independent tests of these authors' hypotheses.

Our underlying aim in this project is to connect Marxian economics with contemporary work on complex self-organising economic systems, a programme often referred to as 'econophysics' (although we will argue in Chapter Five that Marx's own outlook was wholeheartedly probabilistic, and that he can thus be claimed to be just as much a pioneer in this field as he is acknowledged to be in others, such as input-output analysis).

Although the two works just cited both treat the rate of profit as a random variable they do so in rather different ways. In Gibrat the randomness enters as cumulative chance deviation from otherwise identical changes in the value of some variable common to a population. In Farjoun and Machover the randomness is akin to that assumed by the treatment of an ideal gas in classical statistical mechanics. In this analogy the individual value of some variable results from chance interactions between members of the population, even though each outcome is (in principle) deterministic.

Previous work in this area has been confined to studies of the dispersion of *industry* profit rates, among which Glick (1985) has been influential. Not only does our UK company accounts data set permit us to replicate Glick's approach, but we are able to apply his methods at a level of detail which his data did not allow. The result will show that industry-level studies can provide only limited insight into profit rate equalisation.

We also provide a new review of the issues involved in using accounting data to estimate rates of return, and make what we believe to be the first application to economic data of an innovative statistical procedure, the method of *L*-moments, originally developed in the hydrology literature.

The first section of this chapter reviews the significance of the rate of profit in economic theory; the second explains and justifies its treatment as a random variable; the third discusses some testable implications of this approach and some issues in operationalising these tests; the fourth sets out the plan to be followed in succeeding chapters.

1.1 The importance of the rate of profit

The rate of profit is a key concept in classical political economy – by which we mean the tradition of Smith, Ricardo and Marx. On the one hand, individual profit rates are taken to be the motive which leads capitalists to direct investment to one sector or another, and to engage in technical innovation; on the other, consideration of the temporal evolution of profit rates in general leads to theories of stagnation, periodic crisis or irrevocable breakdown. In short, the rate of profit is central to understanding the capitalist market economy as a dynamic entity evolving in time.

In describing political economy as the ‘tradition’ of Smith, Ricardo and Marx, we choose our words carefully. The varying social and political conclusions which can be drawn from the works of these three are well known. But for present purposes what is more important is the theoretical development from one writer to another, and in particular that from Ricardo to Marx, who claimed to solve the conundrum in value theory which defeated his predecessor: namely, how can labour values be said to determine prices even in the presence of capitals having different proportions of variable and constant capital?¹ Marx’s solution to this transformation problem has been controversial since its publication.

There is a long-standing and dominant tradition in economic thought – common to various schools of thought that are otherwise radically opposed – that profit-rates in a

¹ In Volume III of *Capital* Marx remarks that Ricardo ‘certainly feels that his prices of production depart from the value of commodities’ (Chapter 10, 1981:280), although a little later (page 300) he also says: ‘price of production ... is in fact the same thing that Adam Smith calls “natural price”, Ricardo calls “price of production” or “cost of production”, and the Physiocrats call “*prix nécessaire*” ... [t]hough *none of these people explained the difference between price of production and value*’ (emphasis added). Lastly, we read that Ricardo ‘did not understand the adjustment of values to production prices’ (footnote 34, 1981: 305).

capitalist economy have a long-run tendency to uniformity. The best-known version of this is the neo-classical doctrine that holds that all firms will make only a ‘normal’ profit in long-run equilibrium. But a number of writers, including Glick, adhere to a variant of this, which they attribute to Marx: that it is industry, not company, profit-rates that tend to equalise. The transformation problem, according to this school, can be stated thus: given capitals of different organic composition, *and assuming a tendency to profit-rate equalisation between industries*, show how values determine prices.

An alternative conception, proposed by Farjoun and Machover (1983) as truer to Marx’s views, holds that a capitalist economy is normally close to a short-run dynamic equilibrium; in contrast to Glick their interpretation concentrates on profit-rates of units at lower levels of aggregation than that of industries.² These do not equalise, but are instead governed by a non-degenerate distribution whose parameters vary only slowly through time; of particular interest for the present study is their conjecture that company profit rates – *weighted by the capital invested* – should have a gamma distribution. The reason for weighting the profit rates is that they want to make statements about the proportion of the total capital invested in the economy which achieves a given range of profit rates (pages 63 and 64).³

We thus have two different lines of defence of Marx’s solution. Farjoun and Machover contend that requiring *company* profit rate equalisation as a condition of the solution is misconceived. Glick wants to show that the dispersion of *industry* profit rates is ‘narrow’ (in some appropriate sense), albeit persistent. There is no *a priori* reason why these claims should not both be supported by the data. However, they stem from different interpretations of Marx, and a key theme of our work is to provide arguments as to why Farjoun and Machover’s should be preferred.

² Farjoun and Machover do discuss industry profit rates (pages 176-179): they anticipate Glick (1985) in suggesting that industry average rates will fluctuate in a relatively narrow band around a common centre of gravity.

³ Thus although they describe the relevant sample space as the ‘firm space’, it might be more appropriate to refer to it as the ‘capital space’. Their choice of terminology is related to their strategy for dissolving the transformation problem (see below).

These alternative views are discussed more fully below. For the moment the important points are that the transformation problem is the key to debate over the logical coherence of Marx's value theory, and that empirical confirmation or refutation of this theory requires evidence about the rate of profit.

1.1.1 Which rate of profit?

In investigating the empirical properties of the rate of profit as a resource for assessing Marx's economics, we add to a literature of which the first instance (Steibeling, 1890) pre-dates the publication of Volume III of *Capital*.^{4,5} We will quickly be confronted with two problems. The first is the need to specify precisely what is our definition of the rate of profit.⁶ The second is for what entities should we calculate the rate of profit.

The first question has pre-occupied a number of predecessors, and our notice will be directed in particular to Gillman (1956), Glick (1985) and Moseley (1992), as well as a number of other studies both within and without the marxist tradition.

These writers have discussed a wide range of possible definitions. This range extends from those measures which take profit to be the money surplus arising directly from production (Gillman dubs these 'marxian' measures) to those which take it to be simply the net revenue potentially available for distribution to shareholders ('capitalist' measures, in Gillman).⁷ For reasons which differ according to the nature of the question each writer is

⁴ Engels gave Steibeling credit for effort, but for little else (Engels, 1981 [1895]: 109–111).

⁵ This literature is sometimes referred to as 'quantitative Marxism', after the title of Dunne (1991). But much of the theoretical discussion of the transformation problem is conducted by way of algebraic analysis or by arbitrary arithmetic examples, demonstrating that it is possible to be quantitative without being empirical. The reverse is also possible (some evidence is essentially qualitative). As reference to Dunne's introduction makes clear, it is the intersection of the two sets which his contributors have in view.

⁶ As any worker in this field must be: see Ashley (1910) for an early counterpart of our review. Among the sources he discusses are some relied on by Gibrat in his pioneer work on profit-rate distributions.

⁷ Some, but by no means all, of these differing notions of profit are matched by definitions of capital which are acceptable in terms of the logical relations between capital and net revenue.

attempting to answer, measures towards the ‘capitalist’ end of the spectrum tend to be preferred. We will argue below that, at least for our central purpose of testing Farjoun and Machover, Gillman’s marxian measures are more appropriate. Our reasons for this include our assessment of a further issue, that of whether or not accounting data is intrinsically suited to calculation of the rate of return. Accounting data, it will be claimed, is not merely adequate to this task, but essential to it.

Our second question concerns the entities for which should we calculate our chosen measure. Most of the writers mentioned above are concerned solely with the economy-wide rate of profit (which is of course the average rate per unit of money capital). Leaving theoretical considerations to one side and considering solely the numerical aspect, for them the choice of measure affects only the level of the profit rate calculated. However it is intrinsic to the literature to which we are contributing that one must compare profit rates in sub-units of the whole economy – of industries, in the case of Glick; of firms, in the greater part of the present work.

Glick considers only industry average profit rates and their dispersion, but we will argue that questions about their distribution require consideration of higher moments of the data than the first and second. As we will show, not only the higher moments but also other important statistical properties of the profit-rate distribution vary greatly according to one’s choice of its measure. Further, Glick’s neglect of intra-industry distributions vitiates his procedure.

In contrast to Glick, Farjoun and Machover’s hypothesis is not about the distribution of profit rates across industries, or even across firms, but across the total capital advanced (Gibrat’s hypothesis appears to apply to firms). We will thus be investigating both firm-level and capital-level distributions.

1.2 The rate of profit as a random variable

The notion of treating the rate of profit as a random variable has not attracted much interest before now.⁸ Nonetheless, recent developments in economics as a whole suggest that it is one whose time has come.

These are reviewed in a recent Feature in *The Economic Journal* on markets as complex adaptive systems. Its objective is to throw light on ‘a whole swathe of economic phenomena’, *including persistent heterogeneity*, ‘that are anomalies in mainstream economics’ and thus call into question ‘[t]he applicability or not of [its] optimisation framework’ (Markose, 2005: F159). As one of the other contributors writes, ‘A more conventional title for [my] paper⁹ would have been “Markets and Price Formation” but I feel there is a subtle bias in this terminology. For “price formation” seems to presuppose that a single price characterises most exchange activity, and this is precisely what I wish to deny’ (Axtell, 2005: F195, footnote 191).¹⁰

Among other things the Feature investigates situations in which the dynamics are non-computable and solutions can only self-organise through adaptive or emergent processes. These can explain the power law distributions found empirically for many economic quantities, such as income and wealth. Demonstrating these for the profit rate measures we will be examining is one of the contributions of the present work.

⁸ To examine why this is so would be well beyond the scope of this project. But part of the answer may lie in a traditional prejudice in favour of determinism on the part of many Marxists (but, emphatically, not on the part of Marx himself); this would help to explain the neglect of Farjoun and Machover in the Marxian literature. And another part of the explanation may be the ambiguous status of the rate of profit as a concept in neo-classical economics, divided into ‘normal profit’, assumed to equalise in the long run, and ‘rent’, composed of whatever is left above the ‘normal’ profit; this could explain the neglect of this aspect of Gibrat’s otherwise extensively-cited work (for a review of Gibrat’s influence on the literature on firm size and growth, see Sutton, 1997).

⁹ Actually titled ‘The complexity of exchange’.

¹⁰ Axtell’s paper in fact compares the computational difficulties associated with different mechanisms for arriving at general equilibrium. But his point is that decentralised exchange mechanisms – for example, Walrasian groping – are less computationally formidable than an Arrow-Debreu auctioneer.

Evidence for such an explanation can be found by computer simulations of the dynamics of large numbers of interacting agents, described as either agent-based computational economics or econophysics, depending on the disciplinary background of the practitioner. A very pertinent example is Wright (2005), who shows that a very simple model embodying stylised features of a capitalist economy can, for a whole range of variables, generate distributions which correspond to those widely accepted as stylised facts about their real-life counterparts.

One quantity for which Wright is *not* able to demonstrate clear correspondence of distributional form is the rate of profit. This is precisely because there has been, until now, no systematic investigation of empirical profit rate distributions. However, the distribution his model produces has qualitative similarities to some of those we shall exhibit later on.

Use of the word ‘econophysics’ to label this field signals that the underlying notions are shared with a quite different field, statistical mechanics. This is the underlying point of view of Farjoun and Machover (1983). Their rejection of uniformity of profit rates as a condition of the problem is motivated by an analogy with the statistical mechanics of an ideal gas. A detailed account of their reasoning is given in Chapter Five. The key point for our study is that their analogy provides a heuristic guide to deriving a possible form for the distribution of the rate of profit.

Discussions of the transformation problem generally conclude that it is motivated by Marx’s desire to show that ‘prices of production’, formed as an outcome of the competitive process, serve to redistribute the value created in production in such a way as to bring about ‘capitalist communism’: a situation where capitalists whose production process involves higher-than-average proportions of fixed capital nonetheless are able to achieve the same rate of profit as those using lower-than-average proportions (who would otherwise, on the assumption that value is created only by labour, achieve higher rates of profit).

In an ideal gas the velocities, hence speeds and kinetic energies, of each of the individual particles which compose it are taken to be both widely differentiated and rapidly changing

as a result of constant interaction between them, in the form of collisions. Conservation of energy requires that if one particle gains energy from a particular collision the other particle must lose it. Thus the particles are in a sense competing for a share of a fixed pool of energy.

It can be shown that in a gas at equilibrium the ‘most chaotic’ partition of total energy among the gas molecules results in a gamma distribution, in the sense of maximising the entropy of the gas (equivalently, minimising the energy available to do work: see Chapter Five). If capitalist competition is a very disorderly mechanism for allocating a fixed pool of surplus value among capitalists, Farjoun and Machover suggest, then perhaps the rate of profit also has a gamma distribution (page 68).

This raises two issues which would take us far beyond the limits of this study if we tried to fully pursue them. The first is that of what kind of equilibrium, and hence – pursuing the gas analogy – what if anything an economic equivalent to entropy might be. To this we respond, with Farjoun and Machover, that it is an open question (page 239).

The second is whether Farjoun and Machover’s approach has any foundation in Marx’s own writing. Our position is that it does – indeed, that Marx’s approach was clearly and consciously statistical from first to last. We can do no more than indicate our case here, but further evidence is presented in Chapter Five.

Briefly, Marx’s discussion of prices of production relies on the idea that they are in some sense formed by the constant fluctuation in market prices in a competitive economy. His first early statement of this – his Notes on Mill (1844) – is intrinsically probabilistic in character. The position set out there is restated in his *Wage labour and capital*, written as lectures for the German Workers’ Association in Brussels in 1847; thus in the same city as Quetelet, and in the aftermath of the latter’s announcement of his concept of the ‘average man’ (Quetelet, 1842 [1969]). By the time Marx was drafting *Capital* Volume III we find him describing equalisation of profit rates as the formation of their probability density function and discussing how variations in its shape will affect the relation of the mean to the

whole; he even considers the effect of censoring some of the data. Finally, and perhaps most significantly, he cites Quetelet in the course of claiming that the value product of the collective working day is a random variable.

1.2.1 The transformation problem and the rate of profit as a random variable

As already noted, this project is fundamentally inspired by Farjoun and Machover's defence of Marx's approach to the transformation problem, and one of its key concerns will be to test their hypothesis about the distribution of the rate of profit. We now locate Farjoun and Machover's work in the transformation debate.

Controversy over the transformation problem has flared up at intervals ever since the first publication of Volume Two of *Capital*, when Engels set the famous 'prize essay competition' – challenging Marx's readers to show, in advance of the impending Volume Three, how determination of values by labour time could be reconciled with the achievement of an equal rate of profit in an economy where the composition of capital varied (Engels, 1978 [1884]: 101–102).

One of these periodic controversies was sparked by Steedman's conclusion that Marx's solution to the transformation problem was flawed, that his value theory was thus an irrelevance to the substantive questions to which Marx's analysis of capitalism was directed (Steedman, 1977: 28), and that attempting to sustain it was a hindrance to further development of that analysis. In particular, Steedman claimed (page 14), that Marx's solution was wrong in that in a competitive capitalist economy the rate of profit is not, in general, given by $S/(C + V)$, where S , C and V are aggregate surplus value, constant capital and variable capital, respectively. Extensive debate followed, to which seminal contributions were collected in a volume edited by New Left Books (1981) and in Mandel and Freeman .

The latter work, devoted entirely to rejoinders to Steedman, contained a pioneering statement (Langston, 1984) of what has since become known as the Temporal Single System Interpretation (TSSI); the author's premature death meant that Langston's article

was edited and completed by Emmanuel Farjoun, who also contributed to the volume in his own right (Farjoun, 1984). This latter article in turn contains, among much other material, a brief notice of Farjoun and Machover (1983). Langston and Farjoun and Machover take very different approaches, but they both turn on rejection of what they claim is an essential part of Steedman's rejection of Marx – namely, the assumptions of a uniform profit rate and uniform prices for each commodity.

The core of the TSSI is rejection of the need, asserted by all critics of Marx's solution, for the prices of commodities, viewed as inputs to a given round of production, to be simultaneously determined with those of the same commodities as outputs of that production period (in other words, it rejects inter-temporal uniformity of prices).¹¹ Farjoun and Machover, as we have seen, reject the need to assume a uniform profit rate (and uniform prices for a given commodity) even at a given point in time (intra-temporal uniformity). Given this, the transformation problem is not *solved*, but *dissolved*.¹²

1.3 Objectives

It has already been said that the underlying aim of this work is to connect Farjoun and Machover's probabilistic reading of Marx with the modern interest in complex self-organising systems. Their proposed interpretation of Marx can be justified by reference to Marx's own work. But justifying a reading by exegesis is not the same thing as demonstrating that that reading is scientifically fruitful. To do this we will test their hypothesis that company profit rates (weighted by size of firm) have a gamma distribution.

¹¹ Marx's error is supposed to be that he overlooks the need to transform labour values into prices in such a way that input and output prices are equal. TSSI writers claim that input prices do not have to, and in general will not equal output prices, since these are prices of different periods; today's outputs can never be inputs to the production process which resulted in them, since that began in the now unreachable past. Freeman (1995) provides a general introduction to TSSI, while Freeman (1996), and Freeman (2004) contain further contributions.

¹² TSSI treats intra-temporal uniform prices and profit rates as special cases, and this is its common ground with Farjoun and Machover's approach. Freeman (2006, personal communication) has said that he regards the Farjoun and Machover perspective as the observable counterpart of Marx's system.

Alongside this we will test Gibrat's alternative hypothesis, derived from quite different considerations, of a log normal distribution.

We will use the same data set to investigate the fruitfulness of Glick's alternative reading of Marx, according to which competition can be thought of as a process of 'gravitation' which results in industry profit rates oscillating close to the overall rate of profit in the economy.

The nature of Glick's conception is such that testing it requires not only cross-sectional data of the kind needed to test Farjoun and Machover and Gibrat, but also longitudinal information (we avoid the term 'time-series data' in order to avoid misleading implications about the use we will make of longitudinal information). Rather than track particular profit rates over time, as might be done in conventional econometric work, we will look at variation in the distributional properties of the data through time.

Since our data set allows us to compute profits rates in accordance with a wide variety of suggested definitions the foregoing work can be done without committing ourselves to any particular definition *a priori*. We will thus test each of Farjoun and Machover, Gibrat, and Glick using not only their own preferred profit rate definitions, but those of rival authors. Thus we will be able to consider the possibility that even if an approach proves unsuccessful in its own terms it may yet be rescued by adopting insights from an alternative perspective.

The conventional method of testing distributional hypotheses is to use one's data to estimate the population parameters on the assumption that the hypothesis is correct, and then test the goodness of fit between the data and the estimated distribution. But doing so provides little guidance as to what to do if the fit proves unsatisfactory. In the present case, if neither the gamma nor the log normal appear to be plausible forms for the profit rate distribution, it would still be of interest to see what other distributions might provide a better description.

We thus have a problem of assessment in two dimensions: correspondence of different profit rate definitions with different theoretical perspectives, and of identification of distributional models for the various profit rates. These problems are linked, in that the conclusions we draw about the adequacy of the different perspectives will be based on the distributional properties not only of the profit rate choices with which they are associated, but on those of their rivals.

Recent developments in statistical theory provide methods to readily compare the distributional characteristics of different sets of data. *L*-moments (Hosking and Wallis, 1997), an alternative to the usual Pearsonian product-moments, have gained attention in the hydrology literature precisely because of their advantages in distinguishing between alternative distributional models. We will explore the contribution they can make to the empirical exploration of profit rate distributions.

1.4 Plan of work

The remainder of this work is organised as follows. In Chapter Two we review previous contributions to the profit rate literature in the Marxian tradition, with particular reference to the issues involved in choosing profit rate definitions for such studies. In this section we also deal with objections to the use of accounting date to measure the rate of return.

Following that we take two chapters to introduce our data set and the use we make of it. Chapter Three begins with a discussion of the practical issues in calculating profit rates from our data set, continues with exploratory data analysis, and concludes by pointing out the source of some problems revealed by the exploratory investigation. Chapter Four discusses the use of *L*-moments and size-weighted sampling methods as possible methods for overcoming the difficulties revealed in Chapter Three.

Chapter Five applies these methods to testing Farjoun and Machover and Gibrat by identifying distributional models for a number of profit rate definitions, and considers the implications of the results.

In Chapter Six we first test the notion of profit rate gravitation advanced by Glick (1985) by replicating that approach using our data set; we then apply our own methodology to complete our critique of the gravitation concept.

Finally Chapter Seven summarises our conclusions and suggests directions for further work.

Chapter 2 Measuring the rate of profit

The overall focus of this study is on the empirical distribution of profit rates. In this chapter we consider what definitions of ‘the’ rate of profit should be considered. This involves two separate questions: first, we have to consider whether accounting data can indeed measure the rate of return to investment in an economically-meaningful sense; secondly, given a positive answer to the first question, which of a number of possible definitions of accounting rates of return should be adopted?

The principal contribution of this chapter, therefore, is to review Marxist work on the profit rate and take a view on the issues in dispute. This will be a provisional one; we have the opportunity to explore the distributional properties of a number of profit rate definitions, and thus the evidence of later chapters will be relevant.

Since we intend to test particular distributional hypotheses we will obviously want to use the definitions preferred by the authors concerned, but one must bear in mind the possibility that these may be ill-chosen, in the sense of not properly reflecting the theoretical paradigms within which the relevant work is situated.

It might be, for instance, that profit rates calculated according to Farjoun and Machover’s specification turn out not to have a gamma distribution, as they predict, but that profit rates calculated in some other way do have such a distribution. Similarly, Glick’s preferred profit rate measures provide evidence of gravitation that is different from that afforded by measures which he deprecates.

As it happens, Marxist studies of the profit rate have considered a wide variety of profit measures, and have often concluded that the preferable definition is one quite different in kind to that specified by Farjoun and Machover.

Further, Farjoun and Machover’s proposal is not the only hypothesis about the shape of the profit rate distribution: there is also Gibrat’s hypothesis to be tested, and there are other works which, although not explicitly concerned with the functional form of this variable’s

distribution, do provide some information on empirical densities. The latter, as also Gibrat, are *not* within the Marxist tradition, and some work with quite different conceptions of profit, and hence measurements of its rate, and these ideas should thus also be considered.

The structure of this chapter is thus as follows: first we consider whether accounting data can measure a meaningful conception of the rate of profit; secondly, we briefly survey relevant profit rate studies before reviewing the issues involved in choices of profit-rate measure.

2.1 Can profit rates be calculated from accounting data?

A well-known contribution by Harcourt (1969) has been widely cited in support of the claim that accounting data should *not* be used to calculate the rate of return. Since the rate of return in question is that prescribed by neo-classical economics it might seem surprising that Marxist authors should feel the need to rebut this claim, but in fact several have felt the need to do so. In this section we explain the issues and argue that most of the proposed Marxist rebuttals are beside the point.

The question arises as a result of the insistence of neo-classical economics on subjective evaluation; thus the rate of return must measure the *expected* flow of returns to a given investment. The economic, or internal, rate of return (IRR) is the discount rate which, when applied to the flow of expected revenues resulting from an investment, makes the net present value of the project equal to zero. Put another way, the IRR provides a measure for choosing among investments on the basis that the most preferred one is that which maximises expected income while leaving capital intact.

Harcourt showed that the IRR cannot in general be measured by the accounting rate of return (ARR) even in an ideal world where there is no uncertainty, and thus expectations

are always fulfilled and the rate of profit has an unambiguous meaning.¹³ Harcourt's work was later independently reproduced by Fisher and McGowan (1983). If we follow the latter, at any rate, this result appears to rule out *ab initio* our project of using company accounts data to investigate profit rate distributions. This is because they argue that

‘[I]t is clear that it is the economic rate of return that is equalized within an industry in long-run industry competitive equilibrium and (after risk adjustment) equalized everywhere in a competitive economy in long-run equilibrium. It is an economic rate of return (after risk adjustment) above the cost of capital that promotes expansion under competition and is produced by output restriction under monopoly. Thus, the economic rate of return is the only correct measure of the profit rate for purposes of economic analysis.’ (Fisher, Franklin M. and McGowan, 1983: 82)

Their notion of equilibrium, being long-run uniformity of profit rates, is precisely what we are contesting. If accounting data cannot reliably be used to compute profit rates then our project appears to be stymied. Our answer to this has several strands. In the first place we will draw attention to work that suggests that the problem to which Harcourt-Fisher-McGowan (HFM) draw attention may not be of great empirical significance, for at least some purposes. Secondly, we contest the notion that the IRR is the only economically-meaningful profit rate. Thirdly, we present arguments to show that traditional accounting practices, and the data that supports them, are in fact *more* relevant both for our theory and for the practice of capitalists.

2.1.1 Empirical relevance of IRR versus ARR

An explicit response to Fisher and McGowan, albeit at the level of aggregate data, comes from Dumenil and Lévy (1987b). They estimate various versions of both the IRR (the

¹³ Harcourt asks whether ‘the answer obtained by using the accountant’s measure of the rate of profit [will] correspond with what is known, under the assumed conditions, to be the right answer, namely, that the *ex post* rate of return equals the *ex ante* one’. He shows that accounting measures will be biased to a degree dependent on the pattern of returns from individual machines, the method of depreciation, growth or otherwise of the capital stock, and whether or not financial assets are included in the stock of capital (Harcourt, 1969: 311).

expected rate of return on new investment) and the ARR (the rate of return on the total stock of investments) for the U.S. economy from 1900-1900 and show first, that different versions of the two concepts differ in level rather than time-pattern, and second that the time trends of the two concepts are similar (their estimated IRR is, in effect, a smoothed version of the ARR). Hence, they assert that the HFM problem is not an empirical one.

Two qualifications are necessary: first, the profit rates they compute use *gross* product as the income measure (in other words they do not include depreciation); second, a demonstration that estimates of the general rate of profit will be broadly similar regardless of whether one uses the IRR or ARR says nothing about possible differences in the distribution across firms.

A partial answer to the point about distributions of profit rates is provided by Fritsche and Dugan (1997), who use simulation methods to test how well various measures approximate the IRR. They found that the Spearman rank correlation of company profit rates measured by IRR and ARR was over 0.95, significant at $p \leq 0.01$. The ARR performed better than all variants of a suggested IRR surrogate, the conditional estimate of the rate of return (CIRR), suggested by Salamon (1982, 1985).¹⁴ However, there was strong evidence that both ARR and CIRR were biased estimators of the IRR.

An important point about Fritsche and Dugan's work is that it is addressed to the measurement of the profitability of firms with a range of investments, rather than of single projects; one of the criticisms levelled at Fisher and McGowan was precisely that their findings assumed 'oilfield production' – that firms were simply collections of discrete projects (Martin, 1984: 503–504).¹⁵

¹⁴ In this connection one may also notice Edwards, Kay and Mayer (1987), who argue that valuing assets in accordance with 'value-to-the-owner' principles makes the ARR comparable to the IRR.

¹⁵ This is a pointed criticism. Fisher was the principal expert witness in economics in IBM's defence during the U.S. Department of Justice's anti-trust prosecution of 1969–1982, and the alleged impossibility of using accounting rates of return to identify the existence of monopoly rates of profit featured in his testimony on

In fact, Harcourt himself questioned the practical relevance of his result. While the strength of his original paper apparently lay in showing that accountants would not be able to measure the IRR under ideal conditions (the Golden Age¹⁶) Harcourt pointed out that ‘[i]n non-“Golden Age” situations, the only way of finding out whether expectations concerning rates of profit have been realized is to ask accountants – or to use their tools’ (Harcourt, 1969: 313).

2.1.2 Concepts of income and capital

The IRR enables one faced with a choice of investments to select the option that maximises their income: the maximum value which they can ‘consume during a week, and still *expect* to be as *well off* at the end of the week as ... at the beginning’, to quote Hicks’ definition (as quoted in the Introduction to Parker and Harcourt, 1969b: 4, their emphasis).¹⁷ The focus here is on the notion of (personal) income as the flow of enjoyments resulting from consumption, subject to maintenance of capital intact.

The most rigorous formulation of this subjective approach is due to Irving Fisher¹⁸:

Income is a series of events ... For each individual only those events which come within the purview of his experience are of direct concern. It is these events – the psychic experiences of the individual mind – which constitute ultimate income for that individual. ... Neither these immediate processes of creation and alteration [*i.e* production: JW] nor the money transactions following them are of significance except as they are the necessary or helpful preliminaries to psychic income – human enjoyment. (Fisher, Irving, 1969)

IBM’s behalf. A whole chapter is devoted to the topic in Fisher’s subsequent book on the case (Fisher, Franklin M. et al., 1983: 219–270); a reply from one of the government economists on the case is DeLamarter (1987: 114–115).

¹⁶ In which accountants would in fact be redundant.

¹⁷ Originally in Hicks’ *Value and capital* (1946, second edition); reproduced in Parker and Harcourt (1969a), where the passage quoted is on page 75.

¹⁸ The quotation is from the opening words of *The theory of interest* (1930) pp 3 ff.

But because this is unobservable Fisher is forced to introduce first ‘real income’:

those final physical events in the *outer* world which give us our inner enjoyments. This real income includes the shelter of a house, the music of a victrola or radio, the eating of food ...’ (Fisher, Irving, 1969: 35, Fisher’s emphasis)

and then, because the elements of real income are incommensurable, to admit the ‘cost of living, a measure of real income’; in other words, he settles on expenditure on consumption for practical purposes.

However, Simons (1969) suggests an alternative notion of personal income: a ‘purely acquisitive concept having to do with the possession and exercise of rights’, where property rights are ‘any mortgage against the community’. Its measurement ‘implies ... measuring the results of individual participation in economic relations for *an assigned interval* and without regard for anything which happened before the beginning of that ... interval *or for what may happen in subsequent periods*’ (Simons, 1969: 68; the first emphasis is Simons’, the second is ours). Hence Simons’ income, or ‘gain’, ‘may be measured *and defined* most easily by positing a dual objective ... consumption and accumulation’ [our emphasis].¹⁹

A similar conception of (*ex post*) income as ‘consumption and accumulation *excluding* windfalls’ is due to Kaldor (1969: 167, our emphasis). This is a firmly anti-subjectivist contribution, which makes fun of Fisher for double-counting (page 163) and of Hicks (and Hayek) for inconsistency: ‘We cannot define income as what is left after maintaining capital intact and then define the latter as what is required to maintain income intact, without getting involved in circular reasoning’ (page 172). One admitted advantage of the expectational approach to income is its exclusion of windfall gains. Kaldor points out that this serves a clear purpose in accountancy: to make it possible for the proprietors of a

¹⁹ More doubtful is Simons’ further comment that it is ‘folly’ to regard income as a flow and, ‘more emphatically’, as a quantity of goods, services, receipts, fruits, *etc.* The references to ‘services’ and ‘fruits’ suggest that his target here is the subjectivist notion of income as a flow of *enjoyments*, and that what he really wants to emphasise is gain as an *addition to a stock* of rights (presumably measured in money).

business to judge whether it is a success in the sense of fulfilling the expectations which led them to invest their capital (we shall consider in a moment who the ‘proprietors’ are) and claims that it is not surprising that the accountant’s definition of income *ex post* is based ‘as it only can be based’ on a series of admittedly arbitrary conventions.

Thus economists, taken to be in search of a notion of income as ‘value generated in the process of production’, are similarly in search of criteria by which the saving corresponding to productive activity devoted to investment can be separated from that which merely reflects revaluation of future prospects, and ‘to a considerable extent’ their interests run parallel to those of accountants.

2.1.3 Capitalists and accounting

The idea that traditional accounting practices aim at achieving what (political) economists are interested in has been argued still more strongly in two contributions from Bryer (1993, 1994). According to Bryer (1993) modern financial reporting (MFR) is founded on the principles of cost-based accrual accounting and the development of MFR in the late nineteenth century was informed by a conscious opposition to economic ideas of income. His argument is directed against those accounting historians who argue that MFR had no clear conceptual basis and that managers were able to manipulate published accounts so as to advance their interests in opposition to those of the shareholders. On the contrary, the purpose of MFR was to make the managers of the newly-prominent joint-stock enterprises accountable to ‘investor capitalism’.

Moreover, Bryer (1994) disputes Steedman’s claim that:

The ‘value rate of profit’, used by Marx, is of no concern to capitalists, it is unknown to capitalists and there is no force acting to make it equal as between industries. (Steedman, 1977: 30)

On the contrary, according to Bryer MFR is precisely how the value rate of profit is not merely known to (managers of) productive capital but is imposed on them as a constraint,

on behalf of capitalists collectively. Steedman's mistake, it is said, is that he fails to grasp Marx's prescient vision of 'social capital', or investor capitalism: for Steedman there are 'iron industry capitalists' and, in general, 'sets of capitalists' (Steedman, 1977: 39), whereas Marx foresaw the transition to '[c]apital [as] ... a social power, with the [individual] capitalist as functionary' (Marx, 1981: 373).

Modern financial theory holds that rational investors should hold the market portfolio – 'a fully diversified, value-weighted combination of all available securities'; to the extent that they do so (and Bryer claims that by and large they do²⁰) *all* investors own *all* firms (page 316).²¹ Marx's general rate of profit (GRP) is the weighted average of the different rates of profit achieved by individual firms (Marx, 1981: 262) and emerges in competition as 'capital withdraws from a sphere with a low rate of profit and wends its way to others that yield a higher profit' (Marx, 1981: 297). The GRP is established in the context of 'total capital', the ultimate 'fully-diversified' portfolio, where 'the movement in one sphere of production will cancel out the movement of another, [and] the forces mutually counteract and paralyse each another' (Bryer, 1994: 316-317, final quotation from Marx, 1981).

Although capital (and especially fixed capital) is not easily shifted from one 'sphere' to another in the short run, profit rates averaged over time are much the same in one branch as in another (Marx, 1981: 311); according to Marx '[c]apital soon learns to reckon with this experience' and adjusts the required rate of return to allow for differences in risk: thus the GRP is not merely a result, but 'an actual *presupposition* of the capitalist mode of production' (Marx, 1981: 275, emphasis added).

Thus the GRP, while it may emerge from competition (indeed, we argue that its level is determined by the competitive process) becomes a required return that must be met if investment is to be undertaken or allowed to continue.

²⁰ On the authority of Griffin (1982: 40).

²¹ In a later article Bryer makes it clear that the objective is stewardship information for social capital, not information about cash flow for individual savers (Bryer, 1999: 683-684).

2.1.4 Summary

Simons and Kaldor reject the subjective view of income (and hence of capital, in view of the ‘as well off at the end of the week’ principle) on the grounds that the important issue is ‘gain’ – accumulation (of capital) and consumption (by the capitalist).

This has obvious similarity to the marxist view of capital as (self-) expanding value with the consumption of rentiers and industrial capitalists alike being a deduction from accumulation.²²

Kaldor only accepts the expectational approach *ex post* as a way of distinguishing windfall gains from the actual results of the intended activity, so that these latter may be compared with the intended results that induced the original investment. This is isomorphic to Bryer’s argument about social capital’s desire to account for the stewardship of managers. And it also reminds us of Harcourt’s own answer to his problem, given in his original contribution: ‘the only way of finding out whether expectations concerning rates of profit have been realized is to ask accountants – or to use their tools’.

We conclude that the use of accounting data is entirely justified, not as a mere feasible expedient, but as the required method of measuring capitalists’ rate of return.

2.2 Issues in the rate of profit

In this section we review issues to be considered in choosing definitions of the profit rate. An overview of some key studies of the profit rate introduces the discussion.

We begin with Marxist studies of the profit rate, since this paradigm is the background to Farjoun and Machover’s claim that the rate of profit should have a gamma distribution, which is the core hypothesis to be tested. A second sub-section reviews non-Marxist studies that have a bearing on the distribution of the profit rate.

²² ‘Accumulate! Accumulate! That is Moses and the prophets!’ (Marx, 1981: 742).

The earliest Marxist studies are concerned with aggregate data since their interest is the evolution of Marx's general rate of profit through time, envisaged as a test of his claim that there is an inherent tendency of the rate of profit to fall.

Gillman (1956) interprets Marx's comments on the falling rate of profit as a prediction of a long-term secular trend (pages 27 and 29) resulting from a similarly long-term rise in the organic composition of capital (OCC), given an assumed constant rate of surplus value. Noting that various commentators have questioned both premises, Gillman tests this by examining seven different measures of the general rate of profit (total profits divided by total capital) of the capitalist economy, along with the associated measures of the OCC and rate of surplus value.

He begins by computing so-called 'flow' rates of profit, one neglecting depreciation for the period 1849–1939 (page 36) and another including depreciation for 1919–1939.²³ It turns out that if these measures show any long-run trend it is *rising* (with the exception of the post-1929 years), not only for the profit rate, but also for the rate of surplus value and the organic composition of capital (charts on pages 39 and 41, and Appendices I and II).

Gillman then turns to stock measures. The results are equally embarrassing, from the point of view of vindicating Marx: although Gillman is able to show a strongly decreasing trend in profit rates from 1880 to 1920 (with accompanying rise in OCC, but also a rise in the rate of surplus value), this trend disappears for the period 1920–1952; in the case of Gillman 4 it even shows a slight reversal.

Thus the last three measures tested by Gillman are intended to explore the possibility that 'new conditions of capitalist production' mean that Marx's law should be 'reformulated' (page 66ff).

²³ His source is the U.S. Bureau of the Census data for manufacturing industries.

These new conditions are, firstly, the rise of monopoly-capital ('industrial and banking combinations' aimed at reducing competition, controlling investment and output, and eliminating price-cutting practices); second, the advent of scientific management and other developments which increase productivity with less-than-proportional additions to constant capital; third, increases in the cost of realising the surplus value produced ('increasing cost of doing business'). Taking these into account enables Gillman to demonstrate a modest falling trend to the profit rate over the period 1919-1939 (for Gillman 7, 1919-1949).

Moseley (1991): In the wake of Gillman's work it became customary to assume that the rate of profit to be studied was that corresponding to the perception of the capitalist, in other words after both deduction of unproductive expenditures and distributions of surplus value to other classes of agent. This is exemplified by a series of studies of the general rate of profit of the U.S. capitalist sector, of which Mage (1963), Weisskopf (1979, 1981), Wolff (1979, 1986) and Moseley (1991) are particularly notable, Moseley's work being a survey of, and reply to, his predecessors.

This work is a response to a series of attempts to track the long-term evolution of the general rate of profit of the U.S. capitalist sector. Moseley says Marx is 'not always clear' about the precise time-period to which the law of falling profit rate is meant to apply – long-run secular trends, medium-run long waves, or short-run cycles (page 1) – but his interpretation is that it applies 'first and foremost' to long-wave expansions of 30 to 40 years.

Chapter Two provides a tabular summary of the approaches taken by his predecessors to various conceptual issues affecting the measurement of the profit rate, which we adopt, including its notes. These conceptual issues are discussed in turn below.

Moseley, like Gillman, concludes that the increasing weight of unproductive expenditure is the source of declining overall profitability, and part of his contribution is his particular answers to the various problems raised by this issue. However, a key difference between

Moseley and at least some of his predecessors is over whether capital should be measured in money, or in the labour contained in commodities purchased by capital.

The importance of this last point is that it opens the way to rejection of the core of the traditional transformation problem, the supposed need to not merely to show how labour values govern prices, given unequal compositions of capital, but to have the values and prices of inputs equal to those of the outputs (although Moseley does not take this as far as adherents of the temporal single system interpretation described in Chapter One).

Table 2.1: conceptual issues in Moseley (1991)

	Moseley	Wolff	Weisskopf	Gillman	Mage
Labour or money	Money	Labour	Labour	Money ¹	Both ²
Non-capitalist production	No	Yes	No	No	No
Distinguish non-production capital	Yes	No	No	Yes ³	Yes ⁴
Include residential housing	No	No	Yes	No	No
Taxes on wages	Variable capital	Surplus value	Variable capital	Variable capital	Neither

Source: Moseley (1991:43)

Notes

1) Weisskopf's interpretation is that Marx's concepts most rigorously refer to observable quantities of labour, but that estimates in terms of money are nonetheless reliable approximations; thus his estimates are in terms of money.

2) Mage presents two sets of estimates, one in units of current prices and one in units of labour-hours, but argues that the latter are more rigorously correct.

3) Gillman distinguishes between productive and unproductive capital invested in labour-power, but not in means of production.

4) Mage also distinguishes between productive and unproductive capital invested in labour-power, but not in means of production. In addition, Mage considers the wages of unproductive labour to be a part of constant capital, rather than surplus value.

Sherman (1968) discusses profit rates in connection with issues of industrial concentration and the business cycle. He wishes to examine the expected return to additional investments, and uses the rate of profit on the stockholder's equity for most measures of performance, but acknowledges (page 25) that 'micro-economic theory ... generally considers the total return (including interest) on total capital (including borrowed capital)'. He discusses the distribution of profit rates within size classes of firms, and actually exhibits a density function (pages 110–113) and considers differences in variation within different size classes as well as differences in their means (pages 113–120).

Farjoun and Machover (1983) contribute to the literature on the transformation problem and are thus necessarily concerned with the profit rates of individual capitals. Thus the issue of whether the profits to be counted in the numerator should be gross or net of surplus value paid to claimants outside the firm is posed immediately.

A full review of their work will be given in Chapter Five, where we will see that most of their work is theoretical; they present only two very small-scale illustrations to support their hypothesis of a gamma distribution for the profit rate.

We note, for future reference in Chapters Five and Six, that they also point out (page 62) that a 'firm', theoretically speaking, is likely not to coincide with any actual firm, since these are usually involved in more than one line of business and thus may meaningfully be divided into separate accounting units (whether or not this is actually done in practice).

Glick (1985) aims to substantiate a theory of the competitive process which he attributes to classical political economy – that is, to Smith, Ricardo and Marx. As noted in Chapter One, this interpretation is the idea of competition as a process of 'gravitation'; although Glick thus considers the distribution of profit rates his approach is essentially of a different kind to that of Farjoun and Machover. We therefore discuss his work more fully in Chapter Six, and confine our critique in this chapter to his views on the choice of profit rate measure.

Duménil and Lévy take a rather different approach in their work (1987b, 1990, 1999b). We cited them above as evidence that Fisher and McGowan's complaints about the use of accounting data are irrelevant as a matter of fact (Duménil and Lévy, 1987b). Also relevant to our study is the fact that they are collaborators of Glick and share his vision of competition as gravitation.

Duménil and Lévy (1987) showed that little difference in profit rate trends was apparent when economic and accounting rates of return were compared. In a further paper (Duménil and Lévy, 1999c) looking at both gravitation and trends in profit rates in the U.S. economy they compare 'broad' and 'narrow' definitions of profits, and the corresponding capital concepts (pages 8–11). They conclude that sectoral profit rates do gravitate around a common centre – with the important exception of a group of highly capital-intensive industries with markedly low rates of profit (excluding these industries intensifies an upward trend which they identify post-1982).

Other papers demonstrate how long-term movements in the general rate of profit can be attributed to changes in the type of physical capital associated with different phases of capitalist development, and investigate which industries do and do not display evidence of gravitation in the Glick sense (1994, Duménil and Lévy, 1999a).

The non-marxist studies of profitability are mainly studies of dispersion, without reference to any formal model of the underlying distribution, but we begin with the exception.

Gibrat (1931) is most celebrated for his application of his Law of Proportionate Effect to the question of the distribution of firms by size.²⁴ What is often overlooked is that the largest section of Gibrat's book – 100 out of 300 pages – is in fact devoted to studying the distribution of various forms of income including wages and rents (pages 179–185). He

²⁴ Although according to Mitzenmacher (2003: 235) it appears to have first been announced in a journal article of the previous year (Gibrat, 1930).

envisages broad applicability for the log normal distribution in economic and social studies, including the rate of profit. Thus this hypothesis is the sole known alternative to Farjoun and Machover's.

Stigler (1963) is relevant as it is one of the few studies to explicitly address empirical densities, although it does not address the issue of whether they follow any definite law. Stigler presents histograms of industry profit rates, unweighted by size of industry (Chart 3, pages 38–47) for each of the years 1938–1957. Moreover his focus is the relation between industry profitability, the movement of capital, the dispersion of profit rates and equilibrium (pages 4 and 6): '[W]e ... define ... the dispersion of industry rates of return as the measure of disequilibrium in any year' (page 6). (He argues that dispersion is relatively greater in years of depression.)

Singh and Whittington (1968) also consider the dispersion of industry profit rates, but their focus is on the relation between firm size and the level and variability of the rate of return over time (they also investigate Gibrat's Law, in the form of its implied absence of relation between firm size and growth rate). An important finding is that the variability of profit rates through is smaller in large firms.

2.2.1 Stock or flow measures?

Gillman apparently computes flow measures (ones where the denominator is wages plus constant capital consumed, rather than the fixed capital) largely because they permit him to provide lengthy profit rate series; he notes that they are inappropriate for testing tendencies in the rate of profit because Marx's argument is to do with the capital invested, not the capital consumed (pages 33–34). In what follows, these rates will be referred to as Gillman 1 and Gillman 2 (see also Table 2.2 below, which summarises the profit rate measures to be investigated in later chapters).

Gillman's stock measures are *not* measures of the actual money capital advanced; rather, they are estimates of the current reproduction cost, at current prices, of the stock of physical plant and equipment. This is arrived at by cumulating investment expenditures,

depreciated on the assumptions of 40-year lives for plant and 20-year lives for equipment (page 62). His first stock measure (Gillman 3, in future) excludes, and his second one (Gillman 4) includes, circulating constant capital: the inventory of raw materials, work in progress, and finished goods.

Farjoun and Machover clearly advocate a stock rate, but do not discuss whether inventories should be included in the capital measure; depending on the view taken of this, their profit rate measure is similar to either Gillman 3 or Gillman 4. A precise test of their hypothesis about the profit rate distribution must therefore involve one or other of these definitions. If we assume that the ‘capital assets ... operated by the firm’ include stocks of materials and work in progress mandated by its particular production techniques, and that stocks of finished goods are an outcome of the competitive process, then Gillman 4 must be our choice.

Discussion

Marx is adamant that the essence of capital is that it is ‘self-expanding value’ – see, for example (1981: 21) although his first use of the phrase appears to be – and that the rate of this self-expansion, the profit rate, is the key to capitalism’s development. Since a rate of expansion only seems meaningful if the increment is related to its initial value, it must be the stock rate that is relevant. This is not to say, of course, that other ratios may not shed light on the detailed evolution of the system.

2.2.2 Unproductive labour and capital

Since traditional Marxian measures have failed to demonstrate a secular fall in the profit rate, Gillman argues that ‘new conditions of capitalist production’ – the rise of monopoly-capital, the advent of scientific management, and increases in the cost of realising the surplus value produced – have modified Marx’s law of a falling profit rate (as interpreted by Gillman). Gillman proposes that all three can be captured by considering the rise in unproductive expenditures (wages and other costs) on such activities as marketing, administration not connected with the direct production process, and so on.

Hence Gillman calculates a stock rate of profit with unproductive expenditure deducted from surplus value (Gillman 5) and a corresponding flow rate (Gillman 6) where the unproductive expenditure is also added to the circulating capital advanced. He does this (Gillman, 1956: 101) on the authority of Marx's description of such expenditures as 'a machine which buys or sells the rest of the product' (Marx, 1978: 211).

Glick's tests of his theory use only data from manufacturing industry (thus not even in productive industries as a whole), apparently for pragmatic reasons rather than theoretical ones (see Chapter Six, section 1). But he does not discuss whether a better study would cover other industries, or whether it should extend to services, or whether an extension to services should include ones which are unproductive in Marx's sense.

In line with Gillman, Moseley argues (pages 34ff) that Marx's concepts of constant and variable capital do not include capital invested in either circulation or supervisory activities, even where the latter relate to production workers. Since these activities do not create value, it is inappropriate to consider the capital invested in the equipment they use as transferring value to the commodity. He therefore excludes from his calculations capital invested in (for example) commercial buildings, or furniture or office machines in productive industries, as well as all the capital invested in the finance, insurance and real estate sectors.

Discussion

Moseley's procedure is acceptable in studies of the general rate of profit. But when considering company profit rates one must consider the question of how to account for the activities of firms in unproductive industries as well as the unproductive activities of productive firms.

Although Glick accepts that restriction to manufacturing is undesirable he does not explicitly address the issue of unproductive activities. However, his commitment to definitions of the profit rate which, according to him, are those which guide capitalist's investment decisions (see section 2.3.3) suggests that he might take a similar line to Gillman and Moseley. His silence on this topic is unfortunate, because since his is a study of profit rate equalisation between industries he needs to make it clear whether unproductive industries take part in the equalisation.

For Marx the process of profit rate equalisation involves all industries (1981, Chapter 17 'Commercial capital'). The inclusion of commercial capital in the computation of the general rate of profit tends to lower the average rate of profit and hence prices of production in industry, as he demonstrates with a numerical example.

[S]ince the circulation phase of industrial capital forms just as much a phase in the reproduction process as production does, the capital that functions independently in the circulation process must yield the average profit just as much as the capital that functions in the various branches of production. If commercial capital were to yield a higher average profit than industrial capital, a part of industrial capital would change into commercial capital. If it yielded a lower average profit, the opposite process would take place. No species of capital finds it easier than commercial capital to change its function and designation.

(Marx, 1981: 395)

But we are interested in individual rates of profit, and these are determined by firms' relative success in capturing surplus value in competition with unproductive enterprises

among others. Hence these sectors should be included in the investigation of the profit rate distribution.

Moreover, since capitalists regard all their capital as contributing to their profit, not just that laid out on variable capital, then the capital on which they calculate their profit must include both constant capital invested in production and the capital invested in unproductive activities (which forms part of their ‘grounds for compensation’, in Marx’s words (1981: 310ff), and thus the resulting rate is the one on which investment decisions are made.

Thus in calculating ‘stock’ measures we do not attempt to distinguish unproductive capital in otherwise productive firms, and indeed our data set does not allow us to do this.

There is also the question of accounting for unproductive labour. Although our data includes figures for company employment this relates to total employees, not to productive workers; one might impute the latter figure by calculating sectoral proportions of productive workers from input-output tables, but we have not attempted this. For the profit rate measures we regard as most important the dataset provides a variable called Cost of Sales (see Chapter Three, section 1), which includes all direct costs of production. Although it does not exclude supervisory production workers our judgement is that it is likely to be a better approximation for each firm than sectoral averages.

2.2.3 Interest and taxes (broad versus narrow measures)

As numerous authors point out, uncontroversially, the profit rate concept one uses should be the one relevant to the task at hand. Likewise, it is both true and uncontroversial that the concepts of net income and capital are logically related, albeit hard to disentangle, as our discussion of the accounting data controversy indicated.

Gillman’s interest is ultimately the surplus created in production, and moreover at the level of the capitalist economy as a whole, hence ‘broad’ measures of the profit rate. With

one exception he does not consider ‘narrow’ measures (that is, ones net of business taxes and interest payments).

The exception is deduction of direct taxes from the numerator (Gillman 7), on the grounds that government provides services which help in the creation and circulation of surplus value (pages 103–4). The deduction is of direct taxes because *indirect* taxes have already been deducted as part of his measure of unproductive expenditure, and he claims (pages 101–102) the authority of Engels for the notion that all taxes are deductions from surplus value.²⁵ (Note that Gillman’s claim here is about taxes borne directly by capital; the question of taxes paid by workers, in the first instance, is considered below).

In contrast, Glick (1985) aims to substantiate a theory of the competitive process which requires attention to ‘[t]he logic of a rate of profit [a]s a logic of private investment’ (page 87), thus ‘narrow’ measures. He argues that those profit rate measures are to be preferred which include financial assets in their measure of capital, on the basis that these represent ‘the total cost advanced in order to generate the income of the firm’ (page 88). He specifically deprecates ‘the “marxist” ratio of total surplus value divided by fixed capital’ (page 91). However, he defines and measures a wide variety of profit rate measures.

Glick 1 is the measure just mentioned: capital is taken to be total assets including financial assets, and thus interest has to be included on grounds of consistency. Conceptually it resembles the standard accounting ratio ‘return on capital employed’ (ROCE, hereafter).

²⁵ Gillman’s citation is *Engels on Capital* (Lawrence and Wishart) 1937, unhelpfully without a page reference; I have not been able to identify this source as given, but the British Library has a volume titled *Engels on Capital* (n.d.) London: Lawrence and Wishart, with ‘Printed in the USA Composed and printed by union labor’ on the reverse of the title page; the cover of the copy is also stamped ‘Left Book Club edition Not for sale to the general public’; given Gillman’s location in America, I conclude that this is the edition which he used.

The quotation Gillman has in mind appears to be: ‘In general, it is this unpaid labor which maintains all the non-working members of society. The state and municipal taxes, as far as they affect the capitalist class, are paid from it, as is the rent of the landowners. *Etc.*’ (page 6).

Glick 2 answers an objection due to Duménil (1987a): total assets may include a fictitious component, because inter-industry transactions (*e.g.* in commercial paper) will inflate both the asset and liability sides of the balance sheet, rendering this measure of capital artificial; moreover, differences in the degree of these operations in different industries could distort the measurement of equalisation. Thus this rate measures capital as the firm's net debt plus equity (plant, inventories, and cash). This definition therefore requires exclusion of interest and dividends in the capital market.

Glick 3 is 'a very traditional measure, included to test against 'more justifiable ones': it is profit (excluding interest) divided by total assets.

Glick 4 bears the same relation to Glick 3 as Glick 2 does to Glick 1: it excludes fictitious capital.

Glick 5 is Glick 3, but with depreciation included as part of profit.

Glick 6 is said to be that most frequently used in the empirical literature, lauded as the most accurate guide to investment and thus the measure that should be equalised most strongly: 'This view flows from the vision of the firm as the passive agent of stock holders attempting to maximise the return to equity' (page 89). The book value of equity is taken as a proxy for market-valued equity. Again, this resembles the accounting ratio 'operating return on equity' (ORE).

Glick 7 is the conventional profit margin, profit divided by sales, resembling the accounting ratio 'net profit margin' (NPM).

Finally Glick 8 is intended as an approximation of the 'marxist' ratio of total surplus value divided by fixed capital, which is here taken to mean fixed plant. Glick notes that it is not clear that the marxist view would see this as the rate that would be equalised, even though it is this rate that is often the subject of studies of the long-run trend in the average rate.

Duménil and Lévy (1987b) consider the following profit rates:

(1) the rate of return on investment (RRI), the discount rate which equates the present value of the stream of returns to each year's cohort of new investment with the initial outlay on it, or in other words the internal rate of return on marginal investment. Orthodox theory of course regards the *ex ante*, or expected, RRI as the 'correct' measure of the rate of profit.

The definition of profit for this measure is the surplus before deduction of taxes, interest and depreciation.

(2) the economic rate of return on the whole capital stock (ERRK) for the current period, where profit is the present value of the stream of returns remaining in the capital, using, as the discount rate for each year, the estimated RRI for investments made in that year. With varying technology and prices one would expect the marginal RRI to differ from the ERRK, an average measure.

(3) the accounting rate of return on the whole capital stock (ARRK), calculated using three different depreciation schemes (a) straight-line (b) sum-of-digits (a faster scheme), and (c) 'end-of-life' (the slowest possible scheme, since all depreciation is postponed till the discarding of the asset).

Farjoun and Machover (1983) is part of the literature on the transformation problem and thus necessarily concerned with the profit rates of individual capitals. This immediately poses the question of whether the profits to be counted in the numerator should be gross or net of surplus value paid to claimants outside the firm – that is, should rent, interest and, possibly, taxes be deducted? As discussed in the introduction to this chapter, accounting logic implies that our choice will also determine the definition of capital.

Farjoun and Machover are very specific:

'When interpreting our model in the real world, the capital of a firm must be taken as the total capital assets – valued at present (amortized) prices – operated by the firm, *before*

deduction of any outstanding debts and mortgages owed by the firm on these assets.’ (page 63).

Their definition of profit is consistent with this; it must be:

reckoned after deduction of all production costs, including rent and amortization, but before deducting payments of interest on borrowed capital as well as taxes such as income and profit tax. In other words, just as we take a firm’s capital to include capital owed by it to the banks, so also the interest paid on this borrowed capital is taken to be ... the banks’ share in firm’s profit ... [and] ... taxes ... are the government’s share. (page 64).

Gibrat declares that he intends to leave aside all the difficulties posed in even defining profit (page 180) and instead confine himself to studying its distribution. His examples are taken from sources cited in Perroux (1926 [1996]); Borghet (1898) and Dermietzel (1906) which use the rate of dividends as the measure of company profits. These sources are critically reviewed in Ashley (1910) along with a number of others.

Consideration of dividend rates would seem to place Gibrat outside the scope of our study, which is of the rate of return on capital advanced, not returns to stockholders. However, the point of his book as a whole is to establish the applicability of the law of proportionate effect to virtually every kind of economic data. Given this, and his somewhat cavalier attitude to specifying his concepts, it seems reasonable to suppose that he would have argued that any measure of the rate of return would be log-normally distributed. If so, any measures that do not have a log-normal distribution will constitute a refutation of Gibrat.

Stigler’s capital and profit concepts are quite close to those Gillman describes as Marxian: capital equals total assets, excluding investments in other companies, and ‘all the resources an enterprise requires for the conduct of its activities – cash and accounts receivable as well as inventory, plant, and land’ (page 120). Profit includes returns to both lenders and equity holders (page 3).

Singh and Whittington consider three measures of the rate of return: the pre-tax rate on net assets, the post-tax rate on equity assets, and the dividend return on equity assets, of which the last is not relevant to our study. Their concept of returns includes trading profits and investment and other income, net of depreciation and charges for current liabilities (for example, bank interest) but before tax, long-term interest payments and payments to minority interests in subsidiaries. Their measures are thus intermediate between ‘capitalist’ and ‘marxian’ ones, in Gillman’s terminology.

The first of these profit rate measures is the return on capital employed (ROCE), while the second is an after-tax version of the operating return on equity (ORE).

Discussion

The numerator of the rate of profit is net income, the denominator is the capital advanced to achieve the income, and ‘net income’ means whatever is left over from gross revenue once provision is made for preserving the capital stock. Thus any definition of the numerator logically implies a particular definition of the denominator, and *vice versa*. For example, if I plant corn, the net physical income is the quantity of corn eventually harvested, less an amount equal to that planted in the first place. If I call my profit the amount for which I sell the surplus corn, less the interest on the £10,000 I borrowed to buy a combine harvester, the net capital should include not only the value of the machine but also a debit in respect of the £10,000 I owe the bank.²⁶

But as the previous paragraph’s point about physical and financial assets shows, this still leaves open the question of what – or rather, whose – capital is to be counted, and hence whose income. If we are concerned only with the interests of the capitalist class as a whole then a full and correct accounting for financial assets should net out in both numerator and denominator, leaving the choice irrelevant for the purpose of computing a general rate of

²⁶ We neglect wages and circulating capital, as well as the complications of depreciation.

profit. On the other hand, if we are interested in some particular aspect of the process of capitalist production – for example, the capitalists’ success in extracting surplus labour from productive workers – then we will want to include in the numerator the entire net sales revenue attributed to the product (a broad measure).

Likewise, if we propose, as we do, to investigate the rate of profit enterprise by enterprise, then the question of which enterprise any particular assets should attributed to becomes important. For Farjoun and Machover, whose project is essentially to vindicate Marx’s work in Volume I of *Capital* as against the traditional reading of Volume III, then a broad measure as just described is clearly relevant. Whether or not it equalises, in the sense of having some particular small degree of dispersion, is irrelevant, and therefore there is no need to consider whether the profit rate concept is one that is, or could conceivably be, the object of capitalists’ investment decisions.

For authors such as Glick, who are precisely interested in such decisions, the choice is more momentous. We have serious reservations about his choices here. According to a passage which he quotes (page 10) from Ricardo:

When the demand for silks increases, and that for cloth diminishes, the clothier does not remove with his capital to the silk trade, but he dismisses some of his workmen, he discontinues his demand for the loan from bankers and monied men; while the case of the silk manufacturer is the reverse; he wishes to employ more workmen, and thus his motive for borrowing is increased: he borrows more, and thus capital is transferred from one employment to another, without the necessity of a manufacturer discontinuing his usual occupation.

(Ricardo, 1951: 89)

Glick’s discussion of the ‘logic of private investment’ suggests that he thinks it is the logic of the ‘bankers and monied men’ which is relevant. But the passage from Ricardo on which he relies clearly refers to changes in productive capitalists’ demand for means of

production: it is the silk manufacturer's employment of workmen which is his motive for borrowing.

It implies that the proper measure of capital is not his net assets but instead the total value of means of production which the manufacturer mobilises ('net plant', as Glick calls it). If he borrows £100 from 'bankers or monied men' his net assets do not change: the £100 increase in the manufacturer's current account is exactly counter-balanced by his new debt of £100, and this position does not change if he then draws 100 gold sovereigns from the bank, nor even if he then exchanges the 100 sovereigns for £100-worth of means of production, whether these are bales of cotton or hours of labour-time.

Of course, the manufacturer who finances production with a loan of £100 then has to appropriate a profit at least sufficient to cover the interest on the loan; thus the numerator of the profit rate must be profit before deduction of interest and taxes. But this gives us Glick 8, the marxist measure which Glick says is not a good candidate for equalisation.

Glick prefaces his Ricardo quotation by saying that 'Ricardo argues that much of the movement of capital in the competitive process occurs through the financial system'. We have just argued that while it may well be that the financial system is the medium by which the capital flows from one industry to another, on the face of the quotation it is the logic of the manufacturer, not the banker, which is the motive force. But if the contrary is maintained, that implies the use of Glick 6 (profit/equity), which Glick notes (page 89) to be the one 'most frequently used in the empirical literature', 'lauded as the most accurate guide to investment' and thus the rate which should be equalised most strongly.

As he says, this view 'flows from the vision of the firm as the passive agent of stock holders attempting to maximise the return to equity' or, put another way, it follows the logic of investment by 'bankers and monied men' as opposed to that of manufacturers. As we saw in section 2.1, this is exactly the point of view argued by Bryer in arguing that social capital uses accounting data to monitor the stewardship exercised by managers.

On the other hand, if the logic of ‘manufacturers’ is what counts, then since financial assets not only have transparent rates of return but are readily tradeable, then the only rate of profit that is important is that on the capital advanced for production, as suggested by the quotation from Ricardo.

Glick argues in favour of total net assets on the grounds that they represent the total cost advanced to generate the income of the firm, and since total net assets include total debt consistency demands that net interest should be included in the numerator.

There would seem to be various problems with this:

(i) suppose net interest for a given firm is positive (presumably reflecting positive net assets over and above the value of means of production); this suggests that the firm’s activities include banking or investment as well as manufacturing.

(ii) suppose net interest is negative; this implies net debt to be offset against the value of plant and inventory. Indeed, a firm’s means of production may be entirely financed by borrowing: but then its net assets are zero (there are several cases of this in our own data) and hence its profit rate is either undefined (‘infinite’) or strictly zero.

(iii) in either of these cases, changes in the measure of capital (that is, including financial assets and liabilities) do not capture changes in the quantity of means of production employed.

(iv) since capitalists are supposed to look only at price rates of profit, not value ones, then the competition which is supposed to produce a tendency to profit-rate equalisation must be that among all industries – not just among those in which surplus value is created but including those which appropriate surplus value from the productive sector.

2.2.4 Non-capitalist sector

In macro studies of the profit rate the treatment of non-capitalist activity is relevant, because on one view the resources used by such activity are ultimately deductions from the surplus in production. This holds whether the activity is the luxury consumption of capitalists, or expenditure on health, education and the ‘social wage’ in general.

For example, Moseley describes the government sector as ‘the main form of non-capitalist production’ in the present day (whereas in Marx’s day, he argues, it was household production by servants), though he also acknowledges the production activity of non-profit organisations.

The argument for ignoring this sector is that the money used to purchase labour and means of production is not capital, for two reasons: first, it is not recovered through the sale of commodities, because the product is not sold in the market (in the main, Moseley might have added), but more importantly no money-increment is recovered.

Bryer (2000) argues that the distinguishing mark of the capitalist mentality is that the proprietor accounts for the rate of profit. Non-capitalist enterprises may seek to maximise some concept of surplus, but unless their actions are informed by some regard for an average rate of return on capital their prices are not prices of production, and they cannot be said to be competing for an aliquot share of the total surplus value.

Such enterprises are either missing from our data, if unincorporated, or indistinguishable from capitalist enterprises (except by assumptions to do with size).²⁷

²⁷ A minor exception should be pointed out in the case of companies limited by guarantee, a form often adopted by what are essentially large-scale voluntary organisations which are not charities.

Discussion

Whatever the argument for including non-capitalist activity in macro studies, there can be no justification for attempting to include accounting data for non-capitalist entities in a study of the distribution of the profit rate across capitalist enterprises. This is so whether the entities are non-profit organisations (public or private) or whether they are motivated by gain but not conducted according to capitalist logic (for example, the activities of professional persons such as lawyers or freelance business consultants).

We will see in Chapters Three and Four that the practical problem in this connection is not inclusion, but exclusion.

2.2.5 Money or hours

According to Moseley, Marx's concept of capital is that it is essentially money: part 2 of Volume I of *Capital* is titled 'The transformation of money into capital', while in Chapter Four ('The general formula for capital') capital is defined as money with a characteristic form of circulation, $M \rightarrow C \rightarrow M$, which distinguishes it from money as a means of exchange.

This view is contrasted to writers such as Wolff, influenced by Sraffian interpretations of Marx, who insist that the quantity to be measured is the amount of labour contained in commodities. But Moseley points out that apart from the question of Marx's own view of capital, the labour supposed to be contained in commodities is abstract, social, labour, not actual concrete, specific, labour.

Discussion

In principle one can infer a monetary equivalent of labour (time) from macro data, and this has become an important topic in recent years in connection with the TSSI literature discussed elsewhere. But this is irrelevant to our problem, since all we could do with it

would be to convert money quantities to abstract labour hours at a fixed rate, which would not alter the profit-rate since this is a ratio.

2.2.6 Housing

This is relevant to Moseley (1991) because it is a macro study: he concludes that it should be excluded from constant capital on the ground that at the end of the day housing is a consumption good, ‘not a means of production used to produce commodities for the market’, even when it is rented out by a capitalist enterprise (page 39).

Discussion

For us, if there are such things as capitalist rental concerns they share in the competitive process and we will be including them in our estimates just as manufacturers of foodstuffs will be. To the extent that any firms own residential property for purposes other than producing housing services for sale it is unproductive capital and the arguments in the previous section apply.

2.2.7 Are workers’ taxes variable capital?

Moseley argues (pages 39ff) that the portion of workers’ wages which is taken by taxation is capital advanced for production, and the fact that it is not (directly) income for workers is irrelevant to the capitalist; it still has to be recovered before surplus-value can be appropriated. In the UK context this implies that ‘taxes on employment’, such as employers’ National Insurance contributions, should also count as variable capital.

Wolff treats taxes on wages as surplus value, while Shaikh (1978) goes further and considers how the state divides them between a ‘social wage’ and expenditure for the collective interests of the capitalist class (protection of property rights, *etc.*). Wolff’s version involves viewing these taxes as akin to those claims on total surplus value which take the form of rent and interest, which in turn implies regarding the government as a particular fraction of the capitalist class rather than its ‘executive committee’. Shaikh’s version in effect regards the social wage as an indirect form of variable capital and any remainder from

workers' taxes as a quantity of surplus-value entering into the general rate of profit. Whatever the merits of this from an aggregate point of view, we cannot think of any reasonable principle by which appropriate amounts of these quantities could be imputed to particular firms' accounts.

Discussion

We adopt Moseley's argument, partly because this is forced on us by our data set, which gives total labour costs (including National Insurance contributions and other indirect costs) as opposed to wages, and partly because even if the relevant data were available the alternative is unpalatable theoretically.

2.2.8 Current versus historic cost

Gillman's stock measures are *not* measures of the actual money capital advanced but estimates of the current reproduction cost, at current prices, of the stock of physical plant and equipment. Moseley, however, measures the capital advanced by its current, replacement, cost instead of its historic cost (pages 29–30). This is because 'flow' measures of constant and variable capital are relevant to the price of commodities and the amount of surplus value, while 'stock' measures relate to the rate of profit. Current cost is said to conform to Marx's own concepts of constant capital, variable capital and surplus value, which are defined in terms of sums of money which function as capital.

If the average productivity of labor in the production of means of production increases or decreases, or the value of money increase or decreases, then the replacement cost of the means of production will decrease or increase correspondingly, and so will the current value of the stock and flow of constant capital.²⁸

Very relevantly to our own concerns, Moseley immediately adds

²⁸ Moseley cites Marx (Marx: 200–203 and 427–428, 317–319, 139–141).

In principle these concepts correspond to entries in the income statements and balance sheets of capitalist firms.

Recent controversy on the transformation problem has centred on the so-called ‘temporal single system’ interpretation (TSSI) (see Freeman, 1995, 1998, Freeman and Carchedi, 1996, Freeman *et al.*, 2004); this in effect adopts Moseley’s (1991) contention that the value of the variable capital advanced is simply the money paid in wages and generalises it to all inputs.

Conversion of input prices into values is irrelevant, because any divergence is simply a transfer between capitalists at the end of the previous period of production. Profit rates, and hence prices of production, relate to what happens in the current period. The fact that prices and values at the end of the period may differ not only from each other, but from their respective values at the beginning of the period, is irrelevant. Capitalists cannot go back in time and purchase new inputs at old prices. This temporal approach is contrasted to the traditional ‘simultaneous’ reading of the transformation problem in which values (prices) at different periods have to be equated as mutually determining. In the TSS approach, values at the start of one period determine values and prices at the end of the period. Prices do not determine values in general, though they do modify the value advanced by individual capitalists.

This raises the question of whether current cost data are appropriate, or whether historic cost should be adopted. A hasty reading of TSS texts might perceive contradictory suggestions here. In Freeman (1996:14-15) we read that the logical procedure is to measure the capital stock by the money paid for it:

This is of course what the capitalists do. If I pay £2000 for a computer my advanced capital is £2000, regardless of the computer’s original or subsequent value. It is the value of the money, not the machine, that determines my profit rate. Why should Marx contemplate anything else? His object of study was the self-expansion of *money* capital ...

But further. Consider the endlessly repeated charge: capital stock can fall in value if its

elements get cheaper, restoring the profit rate. Excuse me: suppose the computer which cost me £2000 is now worth £500. How does this make my invested capital equal to £500? I paid £2000. That is what my bank manager wants. That is how my rentiers calculate their returns. It is very unfortunate my computer has depreciated because it forces me to find the lost £1500 from somewhere, but find it I must, or go bankrupt. As for my rate of return, it is a proportion of my advanced capital, that is what I paid in the past, not what my investment is now worth.

This quotation suggests historic cost accounting. On the other hand, Freeman elsewhere (1995) quotes Marx to the effect that cotton bought at 6d/lb transmits 1s/lb to the product when worked up, if the price of cotton has doubled in the interval between purchase and use (see also the discussion in Freeman, 1996:243-9), which suggests current cost accounting.

However, this appearance of contradiction is superficial. The quotation dealing with fixed capital is directed against the notion that the value against which profit rates should be calculated is the labour value of means of production mutually and simultaneously determined with prices, as opposed to a sum of money. The quotation relating to circulating capital is directed against the idea that the value advanced should be measured by quantities obtaining at some point other than the start of the production process (and in particular, by quantities obtaining at the end of the production process).

Neither quotation can be taken as arguing that one should disregard what prudent stewards of capital should do, namely to account for changes in the value of the physical means of production by means of appropriate debits and credits to the capital account.

Discussion

Not all stewards of capital are prudent (or honest). But we take the view that most of them are, most of the time, and that any deviations will be washed out in a large-scale statistical study.

2.3 Conclusion

Our review of profit rate concepts and their measurement has argued that company accounting data is not merely permissible, but *required* for measurement of rates of return relevant to our theoretical concerns. Since the literature on profit rate dispersion and distributions addresses a number of rather different theoretical issues we will want to investigate several different measures.

2.3.1 Use of accounting data

As Bryer, Moseley and Freeman all show, what traditional accounting data measures is the marxist notion of the rate of profit – namely, an objective measure relating new value created to the capital advanced to bring this about. This view is supported, from an entirely different perspective, by Simons' comments on the concept of profit being that of gain. Not only that but, *contra* Fisher and McGowan, the subjectivist school is at least as flawed as any objective concept (Kaldor). However, the notion that the distinction is not practically important (Duménil and Lévy) is an open question at best, since their findings relate to trends, while accounting and economic profit rate distributions at a given point in time could be very different.

2.3.2 Profit rate concepts

Our conclusion about the possibility of an objective measure of the rate of return, ascertainable from accounting data without reference to notions about expectation, is a very general one. It does not help us to select among alternative objective measures.

However, since Farjoun and Machover and Glick represent alternative approaches to the transformation problem we should examine the profit rate measures preferred by each in the context of the others' work.

Our review has suggested that certain measures *must* be considered. First of all, a broad measure corresponding to Gillman's 'marxian' measures is necessary for a proper test of Farjoun and Machover's claim that the rate of profit is gamma distributed. We regard Gillman 4 as the appropriate measure, but Gillman 3 is an arguable alternative.

The alternative, traditional interpretation of the transformation problem underlies Glick's work on gravitation of profit rates. Hence we must investigate the distributional properties of his preferred measures (one which include financial assets in their measure of capital) and contrast the results to those for measures he deprecates. Since some of the measures he considers are in principle identical to traditional accounting ratios we will also test all four of these.

Gibrat's hypothesis about profit rate distributions is not inspired by the classical tradition of political economy; indeed, he is agnostic about notions of the profit rate. However, he is far from agnostic on the question of its distribution. If his law of proportionate effect is as universal as he appears to believe, all profit rate measures should be log normally distributed, and any measure which fails to exhibit a log-normal distribution will constitute evidence against his views. On the other hand, if some but not all measures display log normality in distribution a new field of theoretical investigation would be opened.

Our data set allows us to compute many varieties of profit rate, and hence to evaluate the effects of different choices in the context of different problems. In the following chapter we will describe this data set and the methods by which we construct the profit rates shown in Table 2.1, undertake an exploratory analysis of the empirical distribution of profit rates across firms and show (i) that different types of profit rate measure have distinct empirical distributions across firms, and (ii) that the greater profit-rate variability of small firms, compared to that of large firms, also applies in cross-section.

Table 2.2: profit rate measures to be tested

PRM	Type	Description	Notes
gill.1	Flow	no depreciation	Gillman describes this as the ‘traditional’ Marxist measure
gill.2	Flow	with depreciation	
gill.3	Stock	Fixed capital only	
gill.4	Stock	Fixed and circulating constant capital	
gill.5s	Stock	Fixed capital, diminished <i>s</i> (unproductive labour is deducted from profits)	See also glick.8
<i>gill.5f</i>	<i>Flow</i>	depreciation, diminished <i>s</i>	Not actually calculated by Gillman, but mentioned as a possibility, though he claims that it is ‘less pertinent’ to the practical operation of capitalist enterprise; thus we add <i>s</i> and <i>f</i> to the subscripts
gill.6	Flow	augmented <i>c</i> (and diminished <i>s</i>)	Here Gillman tests the effect of considering unproductive expenditure as a form of circulating constant capital. <i>Note</i> that although his text suggests (page 98) that he intends to augment <i>c</i> instead of diminishing <i>s</i> , it is clear from line 8 of his Table I (page 99) that he in fact calculates it as shown in Table 3.1 in Chapter Three, which does seem the appropriate method; see also lines 13-17 of his Appendix 5, where the full results are reported
gill.7s	Stock	diminished <i>s</i> with taxes	Gillman describes this as the ‘capitalist’ measure; he calculates this for three years only, reported in his Table K, page 102
<i>gill.7f</i>	<i>Flow</i>	diminished <i>s</i> with taxes	Not discussed or calculated by Gillman, but included for comparison with gill.5s and gill.5f
ORE	Stock	Operating return on equity	See also Glick 6
ROCE	Stock	Return on capital employed	See also Glick 1
ATO	Flow	Asset turnover	
NPM	Stock	Net profit margin	See also Glick 7
glick.1	Stock	(profit + net interest)/total assets	See also ROCE
glick.2	Stock	(profit + net interest)/(net plant + inventories + cash)	
glick.3	Stock	Profit/total assets	
glick.4	Stock	Profit/(net plant + inventories + cash)	
glick.5	Stock	(profit + depreciation)/total assets	
glick.6	Stock	Profit/equity	See also ORE
glick.7	Flow	Profit/sales	See also NPM
glick.8	Stock	(profit + net interest + taxes)/net plant	See also gill.5s

Notes: in column 1 profit rates are labelled by the object names used in computations; thus Gillman 1 is labelled gill.1, Glick 1 as glick.1, etc.

Chapter 3 Data and exploratory analysis

This chapter has two main parts: the first discusses the dataset used in our work and the way in which it has been used to calculate the various different profit rate measures introduced in Chapter Two, while the second part presents an exploratory analysis of the empirical distribution of profit rates across firms and draws some preliminary conclusions from this.

This analysis makes two substantive empirical contributions: (i) we demonstrate – for the first time to our knowledge – that different types of profit rate measure have distinct empirical distributions across firms; (ii) we show that the greater profit-rate variability of small firms, compared to that of large firms, long accepted as a stylised fact in longitudinal studies (Singh and Whittington 1968), also applies in cross-section.

The chapter has four sections. In Section 1 we discuss the source of our data and its use in constructing analogues of the profit rate measures discussed in the existing literature; Section 2 displays the empirical density functions at the firm level of 21 profit rate measures derived from the literature, and gives tables of the corresponding summary statistics; Section 3 investigates the very lengthy tails exhibited by all profit rate measures and show that these tails are predominantly comprised of the profit rates of very small firms. In Section 4 we review our findings and discuss their implications for our intended test procedures.

3.1 Data source and construction of profit rate measures

3.1.1 Data source: Financial Analysis Made Easy (FAME)

All data used in this work is based on company accounts data drawn from the FAME database (Financial Analysis Made Easy) published by Bureau van Dijk. The database includes details of over 270,000 companies. Company level data may be used directly to test Gibrat's hypothesis about profit rate distributions, but has to be weighted by the appropriate capital measure to test Farjoun and Machover.

Our data consists of the approximately 108,000 companies included in FAME's so-called Jordan Watch series, which are those with *either* turnover greater than £750,000 *or* pre-tax profit greater than £45,000 *or* shareholders' funds greater than £750,000. A version of this already prepared in the form of *Excel* spreadsheets was made available by Roberto Simonetti, from which we use data for the years 1991-5, allowing us to compile sets of data for all 21 profit rate measures.

However, the items reported vary according to the size of company; only an abridged balance sheet needs to be submitted by a 'small' company, defined as one satisfying two of the following criteria:

- Turnover of less than £2.8 million
- Balance sheet total not exceeding £1.4 million
- Number of employees not exceeding 50

Turnover figures do not have to be disclosed by a 'medium'-sized company, one satisfying two of the following criteria:

- Turnover of less than £1.2 million
- Balance sheet total not exceeding £5.6 million
- Number of employees not exceeding 250

Companies that fulfil certain criteria defined in the 1985 Companies Act are permitted to file without detailed profit and loss accounts. None of these exemptions are available for publicly quoted companies, nor for banking, insurance or shipping companies and any of their subsidiaries regardless of size.

The upshot is that our dataset, which already excludes very small companies, nonetheless has some bias towards larger companies within its coverage; in so far as smaller companies

are included there will be some bias towards companies that are publicly quoted, and towards coverage of the financial services and shipping industries.²⁹ A further implication is that the various profit rate measures will be calculated for varying subsets of the whole set of companies, as discussed in the next section.

A possible alternative source of data might have been the two company accounts data bases offered by the ESRC database located at the University of Essex. These are the Company Accounts Analysis Dataset (CAAD) and the Cambridge/DTI Databank of Company Accounts. The main advantage of these datasets is the long time period for which they are available: for 1948–1990 in the case of the Cambridge/DTI data, and 1977–1990 for the CAAD set.

However, these datasets are based on sampling; in the case of CAAD, stratified by size, defined as the amount of a company's capital, and the sampling ratio ranges from 1:360 for the smallest companies to 1:1 for the top 500; the average number of companies included each year is approximately 3,000. This has the disadvantage of forcing the researcher to accept the compilers' stratification procedure. Even leaving this aside, we wanted in principle to use what the econophysics literature refers to as 'high-frequency' data, although in our case the frequency is not longitudinal (much econophysics literature deals with financial market data recorded hundreds of times per hour), but cross-sectional.

3.1.2 Construction of profit rate measures from FAME

The underlying principle we have followed in constructing our estimates of the various profit rate measures is as follows: identify the conceptual elements going into each definition, and then construct the estimates by using the FAME variables which are the nearest direct equivalent to the concept.

²⁹ In practice this size bias will not be a serious issue, in virtue of the estimation methods to be introduced in Chapter Four.

For example, Gillman's discussion of net income proceeds by considering sales revenue and a sequence of various items deductible from it: direct costs (wages and materials), depreciation, unproductive expenditure, and taxation. In other words, he begins with the total surplus value resulting from production, and subtracts successive portions to arrive at the profit appropriated by the industrial capitalist concerned.

By contrast Glick's discussion begins with profit as perceived by the productive capitalist, and considers whether this should be counted net or gross of interest or taxes.

Table 3.1 lists the FAME items used and describes them.

Table 3.1: definitions of FAME variables

Code	Description	Definitions from Bureau van Dijk
BANK	Bank accounts and cash	included here are cash floats, petty cash and receipts not yet banked. Companies usually have both current and deposit account balances which are also classified as cash
CASS	Current assets	STWI plus cash and short-term debtors and investments
COST	Cost of sales (direct production costs)	will normally include all direct elements of the cost of ordinary activities
DEPR	Depreciation charge	charge to the profit and loss account of the amount by which the assets are said to have depreciated in the year
FASS	Fixed assets	tangible assets plus intangibles and investments
INTE	Interest paid	on long and short term borrowing
LOAN	Loans/overdraft	any borrowings due for repayment within 12 months
OTHE	Other expenses (overheads)	operating costs including goods purchased for resale, other external charges, staff costs, other operating charges, administration costs but not depreciation
PRAT	Profit after tax, or 'net profit'	the profit or loss remaining after tax has been deducted
PRBI	Profit before interest	calculated by adding back the interest paid to the Profit before Tax figure
PRBT	Profit before tax	profits from trading after deducting interest paid but before taxation, extraordinary items, minorities, dividends and other appropriations
SHFU	Shareholders' funds	called-up share capital plus reserves
STWI	Stocks and work in progress	usually raw materials, work in progress and finished goods. Stocks are valued on the basis of their cost and not their sale value to the customer
TASS	Tangible assets	depreciated value of plant and equipment
TAXA	Taxation charge	the tax charge based on the results for the year, include UK corporation tax, overseas tax, deferred tax, ACT write-off
TURN	Turnover (sales)	turnover of goods and services net of all taxes, royalties, investments and other non-trading income

We calculate net income for our Gillman profit rate measures by taking turnover (TURN) and successively deducting COST, DEPR, OTHE and TAXA. For our Glick profit rate measures we use PRAT directly in the relevant measures, and contrast this with PRBI + INTE (in principle equal to PRBT, but see below).

An important result of the different reporting requirements for large and small firms is that there is a bias towards large firms in the data needed to calculate the Gillman profit rate measures; this feature also means that these are calculated for a considerably smaller number of firms than are most other profit rate measures (see Table 3.4 below, and accompanying discussion).

The main difference of principle between PRBI – the broadest measure of net income considered by Glick – and the narrowest Gillman measure is that PRBI includes income from non-production (*n.b.* not ‘non-productive’) activities, such as investment income. We could get from the Gillman measure to PRBI by adding the FAME variable OTHI (other income) and – as we have seen – from PRBI to PRBT by adding INTE.

FAME does not calculate these constituent items unless they are given directly in the company’s return. We would like to gain some insight into how far use of a very large data set renders our procedures robust to procedural, as opposed to substantial differences in constructing particular types of profit rate measure. Some of Glick’s measures are in principle similar to standard accounting ratios (return on capital employed, ROCE, and our interpretation of Glick 1; operating return on equity, ORE and Glick 6; the net profit margin, NPM, and Glick 7). Thus we use PRBT for the net income in our accounting ratios but in the corresponding Glick measures we calculate this as PRBI + INTE. The statistics for these conceptually identical measures are thus in fact calculated for slightly different sets of firms (see Table 3.4).

Nonetheless results are available for between 30,000 and 70,000 firms across the economy as a whole, depending on the particular measure, and these should account for

the vast bulk of the genuinely capitalist corporate sector (this last point will be important in Chapter Four).

The profit rate measures to be examined are derived from three sources. First, from the literature discussed above we compute nine measures based on ones used or discussed by Gillman (denoted ‘gill.*n*’ in Table 2.2 in Chapter Two and in Tables and Figures to follow); these include profit rate measures which attempt to account for unproductive expenditure in various ways, and thus are taken as instantiating the kind of measures also adopted by Moseley. Second, we give four profit rate measures derived from the standard literature on financial accounting ratios. Thirdly, we examine profit rate measures corresponding to the eight measures tested by Glick (‘glick.*n*’ in Table 3.2). Three of the financial ratios – ORE, ROCE and NPM – have very close analogues in Glick’s set of profit rate measures (Glick 6, Glick 3 and Glick 7), differing only in taking the net income measure before and after tax, respectively.

The 21 profit rate measures are combinations of 10 distinct measures of ‘net income’ and eight distinct measures of ‘capital’ (the inverted commas signal that, as will be seen from Tables 3.1 and 3.2, turnover appears in both categories); the definitions of the individual FAME variables are given in Table 3.1, and the definitions of the measures in terms of FAME variables in Table 3.2. (In considering Table 3.2, note that one capital measure appears in both the Gillman and Glick systems, while the three accounting capital measures all appear also in the Glick system; for convenience of comparison within each system these are referred to, in Table 3.2, as ‘k.*xyz*’, where *x*, *y* and *z* are numbers in the system of profit rate measures in which they appear, with the accounting ratios numbered in the Gillman sequence for convenience).

Table 3.2: construction of profit rate measures using FAME variables

profit rate measure	s-measure	k-measure	Construction using FAME variables
Gill.1	s.1	k.1	$(\text{TURN} + \text{COST}) / - \text{COST}$
Gill.2	s.234	k25f7f	$(\text{TURN} + \text{COST} - \text{DEPR}) / (- \text{COST} + \text{DEPR})$
Gill.3	s.234	k.3*	$(\text{TURN} + \text{COST} - \text{DEPR}) / \text{TASS}$
Gill.4	s.234	k.45s7s	$(\text{TURN} + \text{COST} - \text{DEPR}) / (\text{TASS} + \text{STWI})$
Gill.5f	s.5s5f6	k25f7f	$(\text{TURN} + \text{COST} - \text{DEPR} + \text{OTHE}) / (- \text{COST} + \text{DEPR})$
Gill.5s	s.5s5f6	k.45s7s	$(\text{TURN} + \text{COST} - \text{DEPR} + \text{OTHE}) / (\text{TASS} + \text{STWI})$
Gill.6	s.5s5f6	k.gill.6	$(\text{TURN} + \text{COST} - \text{DEPR} + \text{OTHE}) / (- \text{COST} - \text{OTHE})$
gill.7f	s.7s7f	k25f7f	$(\text{TURN} + \text{COST} - \text{DEPR} + \text{OTHE} - \text{TAXA}) / (- \text{COST} + \text{DEPR})$
Gill.7s	s.7s7f	k.45s7s	$(\text{TURN} + \text{COST} - \text{DEPR} + \text{OTHE} - \text{TAXA}) / (\text{TASS} + \text{STWI})$
ROCE (8) ‡	s.8910	k.810 ***	$\text{PRBT} / (\text{FASS} + \text{CASS})$
ORE (9) ‡	s.8910	k.9 **	$\text{PRBT} / \text{SHFU}$
NPM (10) ‡‡	s.8910	k.810 ****	$\text{PRBT} / \text{TURN}$
ATO (11)	s.11 ****	k.11 ***	$\text{TURN} / (\text{FASS} + \text{CASS})$
glick.1 ‡	s.12	k.135 ***	$(\text{PRBI} + \text{INTE}) / (\text{FASS} + \text{CASS})$
glick.2	s.12	k.24	$(\text{PRBI} + \text{INTE}) / (\text{TASS} + \text{STWI} + \text{BANK} + \text{LOAN})$
glick.3	s.3467	k.135 ***	$\text{PRAT} / (\text{FASS} + \text{CASS})$
glick.4	s.3467	k.24	$\text{PRAT} / (\text{TASS} + \text{STWI} + \text{BANK} + \text{LOAN})$
glick.5	s.5	k.135 ***	$(\text{PRAT} + \text{DEPR}) / (\text{FASS} + \text{CASS})$
glick.6 ‡	s.3467	k.glick.6 **	$\text{PRAT} / \text{SHFU}$
glick.7 ‡‡	s.3467	k.7 ****	$\text{PRAT} / \text{TURN}$
glick.8	s.8	k.8 *	$\text{PRBI} / \text{TASS}$

Notes

- (1) Closely analogous profit rate measures are indicated by ‡, ‡‡, ‡‡‡; they differ by taking profit before tax (s.8910, accounting ratios) or after tax (s.3467, Glick measures)
- (2) The following pairs of capital measures are identical: * k.3 and k.8; ** k.9 and k.glick.6; *** k.810 and k.135; **** k.11 and k.7 (which in turn are identical to s.11).
- (3) COST and OTHE are given as negative values in FAME, hence the counter-intuitive construction of the items in which they appear

In Chapter Two we raised, as an objection to definitions of capital that include financial assets and liabilities, the logical possibility of these taking values which were either negative, or strictly zero. A further complication comes from considering that negative capital may be accompanied by either positive or negative net income.

It is far from clear how to treat some of the possible combinations. Consider Figure 3.1, where firms with positive net assets and positive net income lie in the NE quadrant, firms

with negative net assets and negative net income in the SW quadrant, and so on. However, unmodified use of the data to compute profit rates will return positive numbers in *both* quadrants shown shaded in the Figure.

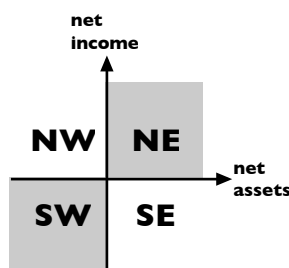


Figure 3.1: income and asset quadrant

Interpretation of the rate of return is straightforward for firms in the NE and SE quadrants. In a profit rate measure that satisfies accounting logic the net income should be counted after payments to the owners of a firm's debt for the use of the funds. Thus the ratio of net income to net capital represents the rate of return to the net value owned by the shareholders (which will be a rate of loss if net income is negative).

However, consider a firm in the NW quadrant, with negative net assets and positive net income; the ratio of the two is negative, which implies loss – but who does this loss fall on? Clearly the firm, and thus its nominal owners, have *gained*, not lost.

On the other hand, in the case of firms in the SW quadrant the net income is a drain on the firm, but the rate of return will be positive, which also seems perverse considered from the point of view of the firm and its owners. But if a firm has negative net assets, then one interpretation is that beneficial ownership in fact lies elsewhere than with the shareholders. Negative net income for the firm implies positive income (above the cost of capital) for the owners of the debt, and the ratio of the two is the rate of return to *their* stake in the firm.

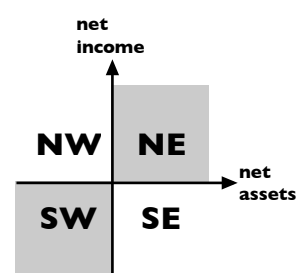
On this basis, the positive net income of firms in the NW quadrant, and the corresponding negative rate of return measures losses to the creditors. But is this a reasonable interpretation, on the assumption that they have been correctly rewarded for the use of their funds, rewards deducted before arriving at net income?

The position we have taken is that net income necessarily implies net value creation, and that recording a negative rate of return when this happens is intrinsically perverse (and that the converse, recording a positive rate of profit when net income is negative, is even more perverse). Since our focus is on firms as entities and their success, or otherwise, and not on owners of money capital, we ensure that firms which have positive (negative) net income have a positive (negative) rate of return by taking the *absolute* value of the capital measure.

That these are not merely theoretical issues is shown by Table 3.3, which records for each profit rate measure (using 1995 data) the proportion of all firms in each quadrant.

Table 3.3: firms with negative net assets

	% in NE	% in SE	% in NW	% in SW	% with total $K \leq 0$
gill.1	98	02	2.0E-03	0	2.0E-03
gill.2	97	03	0	0	0
gill.3	97	03	0	0	0
gill.4	98	02	0	0	0
gill.5s	79	21	0	0	0
gill.5f	78	22	0	0	0
gill.6	78	22	8.6E-02	1.1E-02	9.7E-02
gill.7s	83	17	0	0	0
gill.7f	83	17	0	0	0
ROCE [#]	82	18	0	0	0
ORE [‡]	77	14	3.9	4.6	8.58
NPM ^{##}	81	19	0	3.1E-03	3.1E-03
ATO	100	00	0	0	0
glick.1 [#]	80	20	0	0	0
glick.2	73	16	7.2	4.3	11.5
glick.3	81	19	0	0	0
glick.4	72	16	7.5	4.5	10.2
glick.5	87	13	0	0	0
glick.6 [‡]	77	15	3.9	4.7	8.6
glick.7 ^{##}	80	20	0	3.1E-03	3.1E-03
glick.8	86	14	0	0	0



Notes:

- (1) Table shows the fraction of all observations falling in the quadrants of a chart in which k is on the horizontal axis, and s on the vertical: thus NE gives the number of firms with positive net income and positive capital, using the definitions of the relevant profit rate measure.
- (2) '0' means 'exactly zero'; other entries are to one decimal place or to two significant figures, hence rows may not sum to 1.

As can be seen, nearly 10 per cent of all firms have negative capital under the definition involved in ORE (operating return on equity) and Glick 6, and more than 10 per cent in the case of Glick 2 and 4.

In what follows rates of return are calculated as ratios, rather than percentage rates (in other words, a profit rate of 10 per cent is given in the form 0.1). All data was exported from the original *Excel* spreadsheets to the *S-Plus 2000* statistical program, which was used to perform all calculations, and to produce the charts.

3.2 Empirical densities and summary statistics

3.2.1 Estimating empirical density functions

We begin by considering estimates of the empirical density functions across firms for each of our 21 distinct profit rate measures. Histograms can be shown to be appropriate empirical estimates of density functions (Härdle, 1991); they are unbiased estimators at the mid-point of each bin, and – as intuition suggests – reducing bin width reduces the bias elsewhere at the expense of increasing variance (page 18). However, they suffer the well-known drawback that the density estimate is vulnerable to unfortunate combinations of width and origin for the bins (the bin mesh) (page 26).

Kernel density estimates (KDEs) avoid this problem in the following way: consider the observed values of some sample distributed along the real number line at the appropriate points; at the points representing each observation erect congruent two-dimensional figures (the kernels), scaled such that the total area enclosed by the figures sums to 1; then at each point of the axis sum the heights of the kernels. The result is the KDE. If a differentiable kernel such as the probability density function (PDF) of the Gaussian distribution is chosen, the result will be a strictly smooth estimate; but – since any practicable representation of the estimate will be made using discrete points – triangular or even rectangular kernels will give smooth-looking estimates given enough observations. Even more remarkably, it can be shown (Härdle 1991:76-8) that the choice of kernel – even

between such different forms as the Gaussian and rectangular – is practically irrelevant to the efficiency of the KDE, given appropriate choice of width of the kernel (bandwidth).³⁰

Given the size of our database, histograms of the profit rate measures can have a natural origin at zero, and equally natural bin widths of 0.01 (that is, a one percentage point difference in the rate of profit), without fear of misleading artefacts. This is illustrated for Gillman 1 in Figure 3.2, which consists of a histogram as just described with a kernel density estimate superimposed. Here the scale of the vertical axes is the value of the empirical probability density function, not a count or proportion.

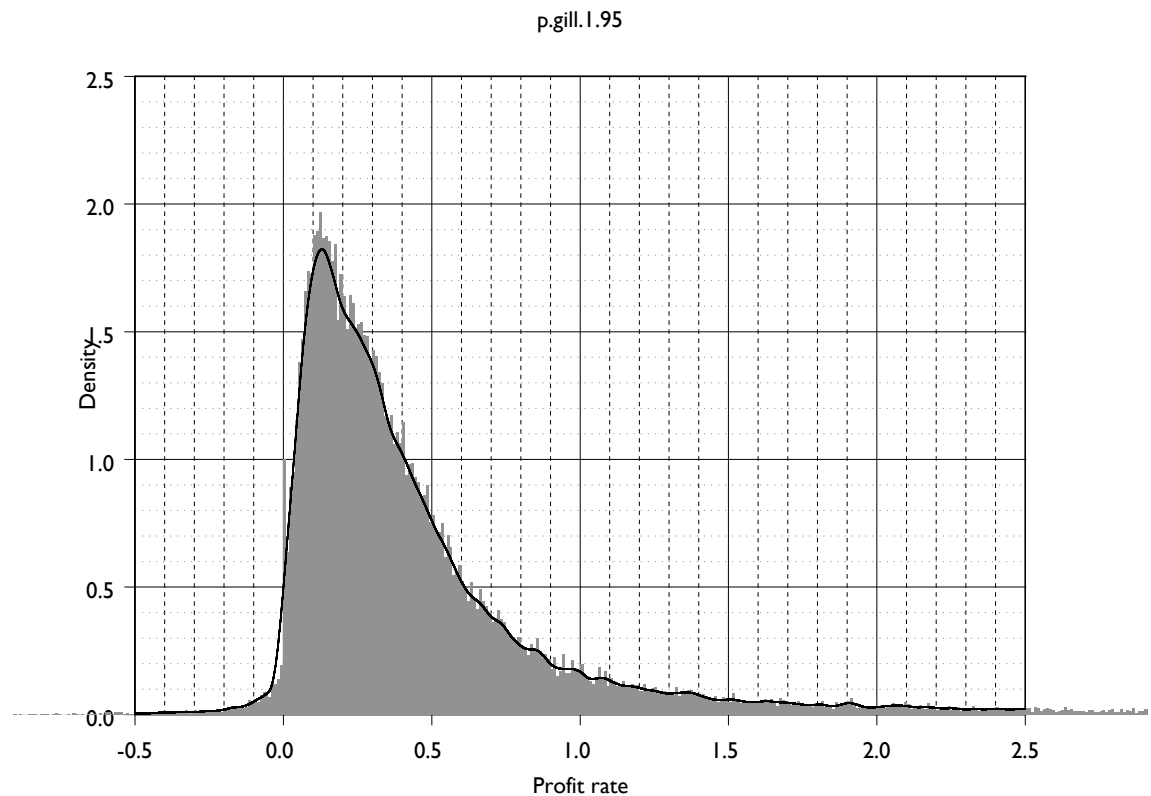


Figure 3.2: empirical density estimates for Gillman I; histogram and kernel methods

³⁰ Nonetheless a cosine-based curve, the Epanechnikov kernel, is theoretically optimal (Härdle, 1991).

Although the narrow bins result in a fairly noisy estimate it is clear that this *is* noise, due to the consequent variance of the estimate: the overall shape of the estimated distribution is in little doubt.

This is emphasised by consideration of the KDE, constructed using a triangular kernel with a bandwidth of 0.1. We choose a kernel with finite empirical support in order to ensure the possibility of density estimates equal to zero in regions with ‘sufficiently’ sparse observations, and a triangular kernel is convenient to implement in the software. The bandwidth is chosen by trial and error, the criteria being adequate penetration of the mode combined with a reasonable degree of smoothing in the tails. Arguably this is somewhat too narrow a bandwidth, as it has failed to smooth out the fluctuations in the right-hand tail; even so, the KDE fails to penetrate the peak quite as fully as might be hoped, and also overestimates the density on the left-hand flank, thus illustrating how the bias of KDEs increases with the curvature of the density function $f(x)$ (Härdle 1991:57).

We conclude that KDEs are not superior to histograms in producing empirical densities for firm-level distributions. Thus in the next section we use the latter to make broad comparisons of the distributions of the different profit rate measures.³¹

3.2.2 Histogram estimates of empirical profit rate densities

Here we present histograms of the 21 profit rate measures to common horizontal and vertical scales (Figures 3.3 to 3.6). Qualitative first impressions suggest some patterns.

We begin by comparing the distributions for Gillman’s two basic ‘flow’ measures, Gillman 1 and 2, with his basic ‘stock’ measures, Gillman 3 and 4 (Figure 3.3). The flow measures differ as to whether depreciation is taken into account, while the stock measures both include depreciation but differ as to whether the capital definition does or does not include stocks and work in progress. We see basic resemblances within each pair, and a

³¹ However, we will use KDEs below to estimate the modes of the profit rate measures.

basic difference between the two pairs; however, subjectively it would seem that the within-type difference is much stronger in the case of the stock measures.

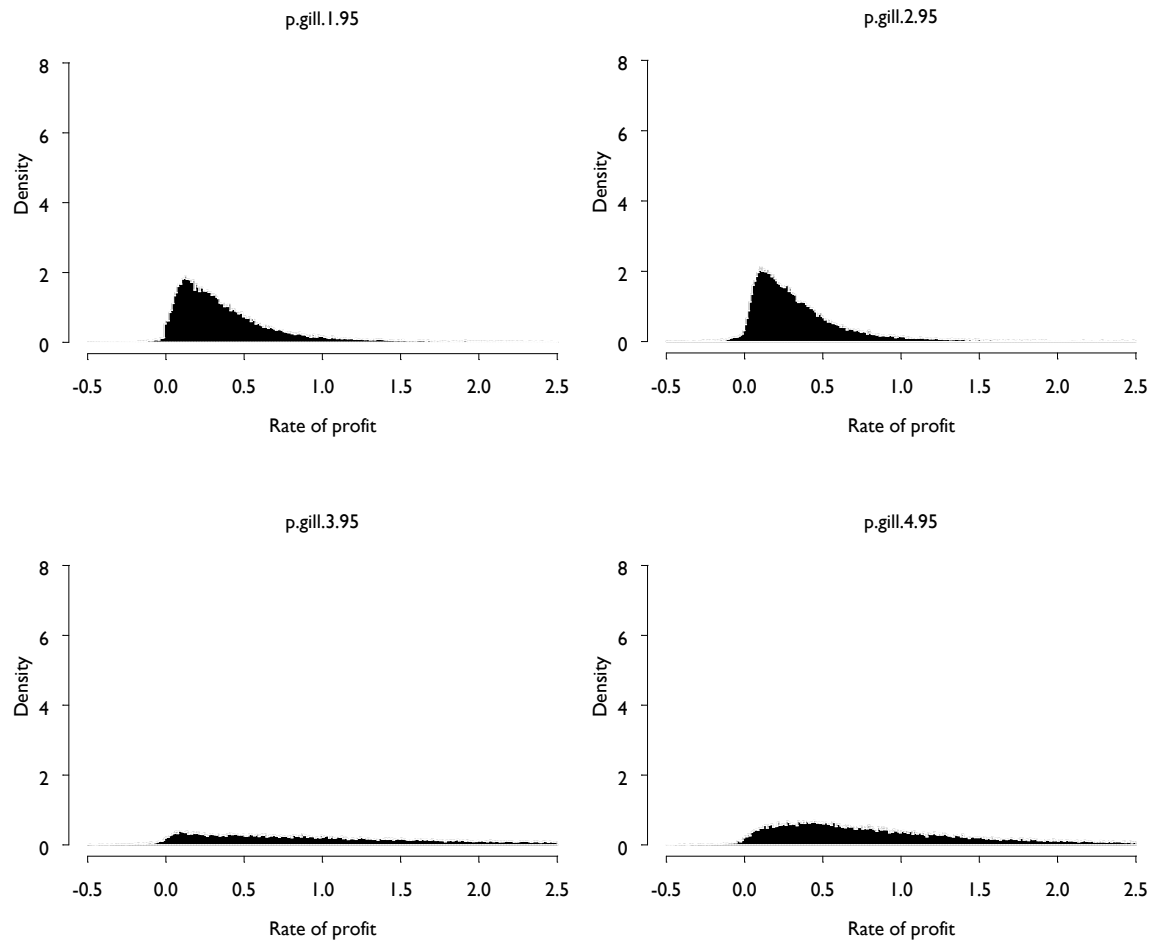


Figure 3.3: empirical density functions, Gillman's 'marxian' measures

The next measures to be considered (Figure 3.4) are the stock and flow versions of Gillman's 'capitalist' measures taking into account, successively, unproductive expenditure (Gillman 5f, 5s and 6) and taxation (Gillman 7f and 7s). Once again, note the similarities between the two stock measures on the one hand, and the three flow measures on the other, regardless of the precise concept of rate of return, which appears to affect only location (and possibly scale), while the stock/flow distinction appears to correspond to a more basic difference in form.

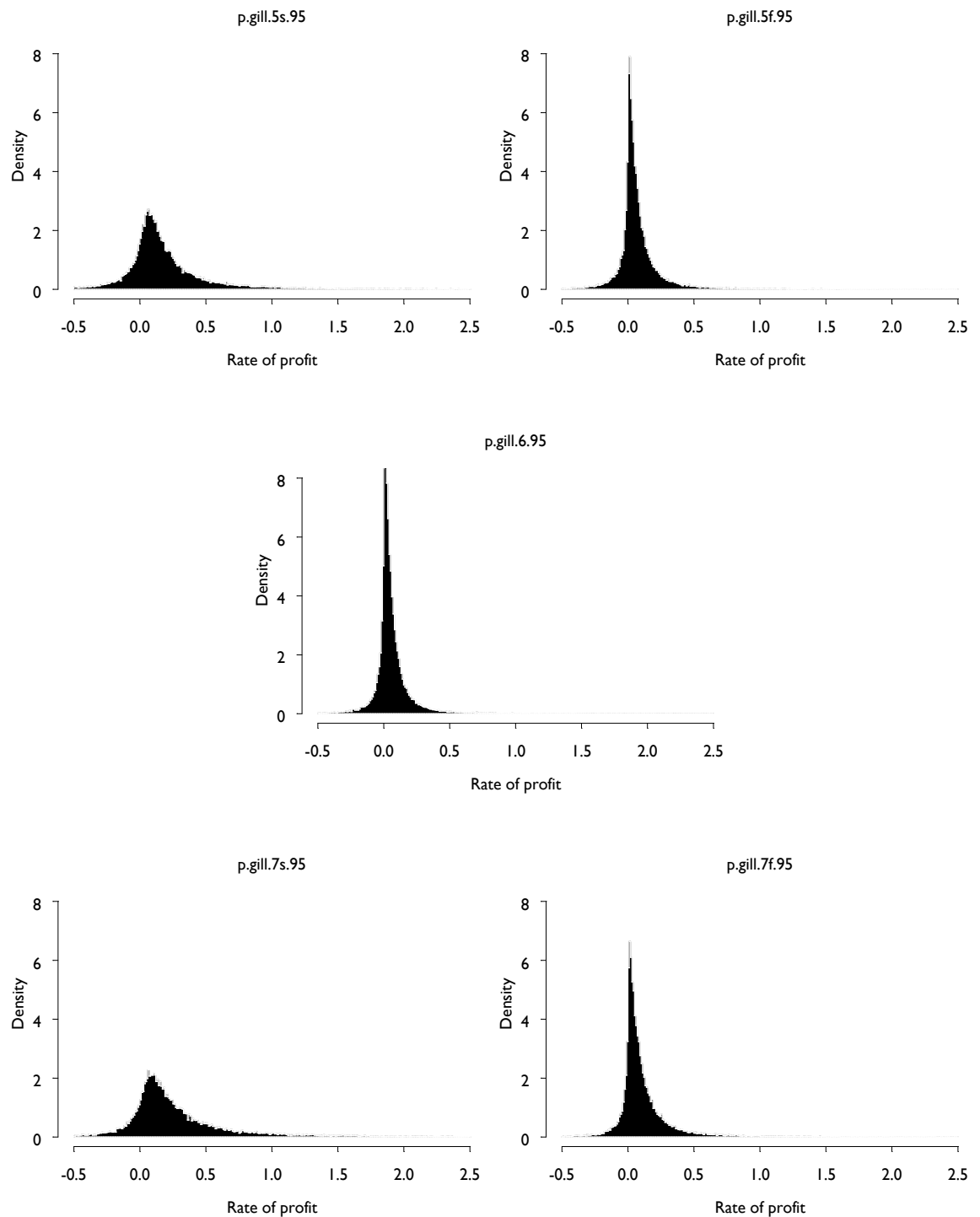


Figure3.4: empirical density functions, Gillman's 'capitalist' measures

We noted above the conceptual links between three of the standard accounting ratios and their counterparts among the profit rate measures considered by Glick: ROCE (\approx Glick 1), NPM (\approx Glick 7), ORE (\approx Glick 6) differ only in whether net income is

counted before or after tax. Not surprisingly the empirical density functions exhibit corresponding similarities (Figures 3.5 and 3.6).

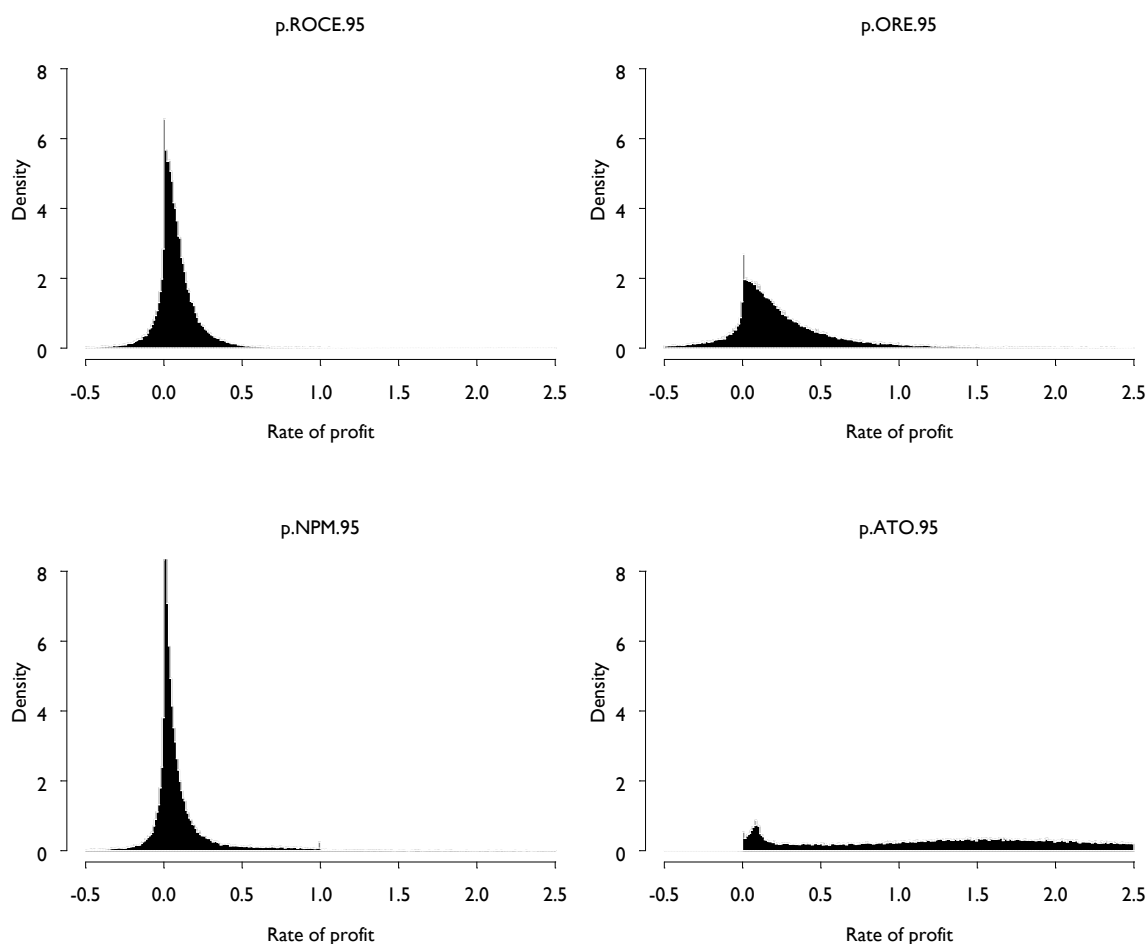


Figure 3.5: empirical density functions, accounting ratios

In fact, all the Glick measures (bar Glick 7) are stock measures, and this suggests that we might be surprised at the systematic differences between the odd- and even-numbered Glick profit rates, especially between the pairs Glick 1/Glick 2 and Glick 3/Glick 4. Since each of these pairs share net income definitions and differ only in the capital definitions (fixed plus current assets/tangible assets plus stocks and work in progress plus bank accounts plus loans), accounting logic suggests that only one of each pair can be acceptable. Since Glick 1 is essentially equivalent to ROCE, the implication is that Glick 2 and Glick 3 should be discounted, since these measures differ from Glick 1 either by numerator (Glick 3) or denominator (Glick 2), but not both.

ROCE (a stock measure) is, as one might perhaps expect, similar to the odd-numbered Glick measures (also all stock concepts). But surprisingly it is also similar to Gillman's flow measures and to NPM (also a flow measure). Finally, note the contrast between all these and Gillman's stock measures 5s and 7s, which are similar to the even-numbered Glick measures.

These patterns are striking, and in themselves demonstrate the possible fruitfulness of a research programme organised around the study of profit rate distributions (in this case, offering further understanding of the accounting issues involved in the rate of return concept).

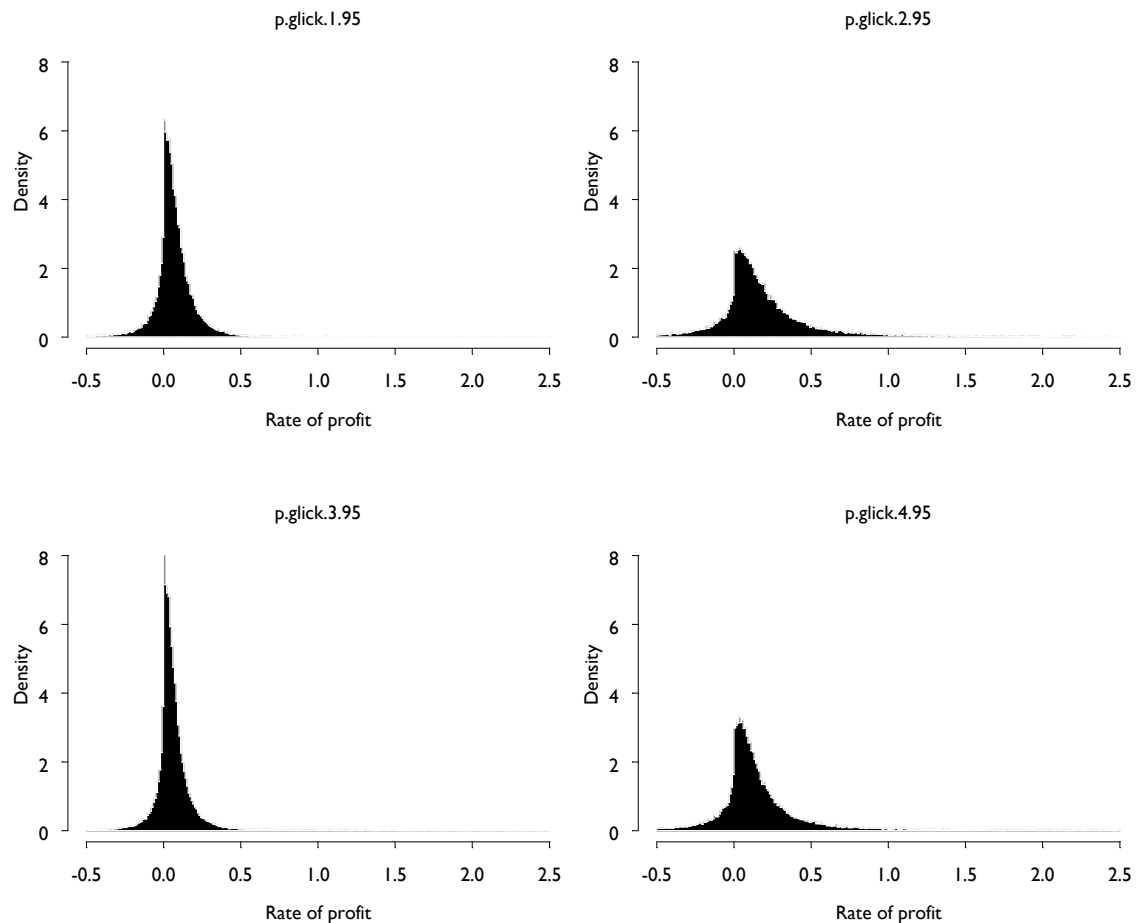


Figure 3.6 (a): empirical density functions, Glick measures

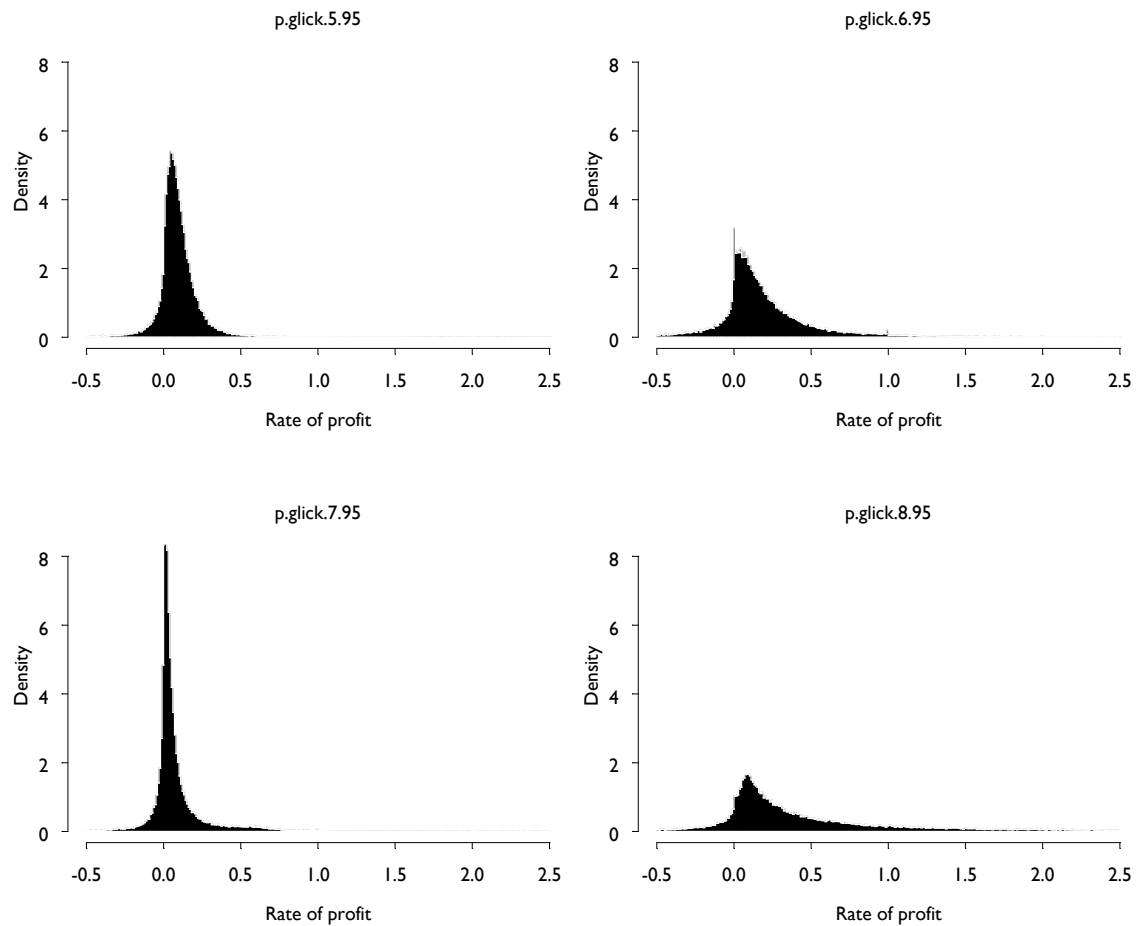


Figure 3.6 (b): empirical density functions, Glick measures

Such explanations would take us beyond the scope of the current project, which is to assess the contribution of Farjoun and Machover to the transformation problem. Nonetheless, these qualitative results suggest some reflections relevant to this, to which we return in the concluding section of this chapter.

One outstanding qualitative feature not noted so far is the exceptional nature of ATO in being clearly bi-modal. We will see in Chapter Six that this can be accounted for as a mixture of radically different industry profit rate distributions. Leaving aside this exceptional case, it is nonetheless possible to discern three broad groups among the profit rate measures.

First, we have two profit rate measures with apparently very high dispersion (large scale), Gillman 3 and Gillman 4; next there is a group of middling scale including such measures

as Gillman 1 or Glick 2; and finally there are measures such as Gillman 5f or Glick 7 distinguished by exceptionally high densities at the mode (one might also remark their possible resemblance to double exponential distributions). This impression is confirmed by reference to the median average deviation (MAD) figures given in the next section (Table 3.4), although not by the other measures of dispersion.

With the exception of ATO, already noted, it might seem that selection and estimation of appropriate distributional forms would be straightforward. In fact, such investigation is complicated by the combination of two facts, neither of which is readily apparent from the charts.

The first is that all the empirical PDFs show definite skewness, although some scrutiny is needed to appreciate this in the case of Gillman 5s to 7f, ROCE, NPM, and the odd-numbered Glick measures.

The second, in all cases, is that although the overwhelming majority of each probability mass is concentrated on a (relatively) small range of profit rates, these masses are flanked by long tails of very low but not negligible density. This feature was evident in Figure 3.2, where we compared KDE and histogram estimates of the densities; a tail of perceptible density clearly stretched out beyond 2.5, the maximum rate of return charted in our histograms, but in the smaller Figure 3.3 this tail apparently peters out around 1.5.

The first group of profit rate measures – our two candidates for testing Farjoun and Machover – clearly have right-hand tails extending well beyond the range of profit rates charted here.³² But note that this is also apparent in the case of Gillman 6, a member of the third group of apparently more concentrated measures (our chart also truncates the peak of this measure's density, which is over 10, as is that of Glick 7). In fact, all the measures have

³² As does ATO, of course.

extreme ranges; this is evident from the conventional summary statistics presented in the next section.

3.2.3 Summary statistics

Tables 3.4 (a) to (c) give summary statistics for each of our three main groups of profit-rate measure.

The first row in each table shows the number of *missing* values, that is, the number of firms for which one or more of the data items needed to compute either s or k was missing; since the total number of firms in the database is approximately $1 \cdot 10^5$ a broad estimate of number of firms for which a given profit rate measure has been calculated can be found by subtraction.

Table 3.4 (a): summary statistics for underlying data, Gillman profit rate measures

	gill.1	gill.2	gill.3	gill.4	gill.5s	gill.5f	gill.6	gill.7s	gill.7f
Missing	5.86e4	6.25e4	6.34e4	6.98e4	6.98e4	6.26e4	6.26e4	7.51e4	6.9e4
GRP	2.2451	0.2162	0.5972	0.4518	0.1232	0.0594	0.0526	0.1844	0.088
Mean	15.7	0.731	126	1.83	0.233	0.154	8.58e17	0.424	0.26
10% TM	0.408	0.346	3.81	1.12	0.159	0.0584	0.0451	0.24	0.0861
Median	0.327	0.286	2.19	0.881	0.121	0.0403	0.0326	0.176	0.0591
Mode.nrd	0.1353	0.1261	0.6746	0.4044	0.0719	0.0143	0.0129	0.0923	0.0204
<i>bandwidth</i>	<i>17.09</i>	<i>14.58</i>	<i>225.79</i>	<i>52.13</i>	<i>10.73</i>	<i>4.12</i>	<i>3.19</i>	<i>14.80</i>	<i>5.57</i>
Mode.alt	0.1322	0.1201	0.0957	0.3918	0.0697	0.0143	0.0129	0.0885	0.0204
<i>bandwidth</i>	<i>10.00</i>	<i>10.00</i>	<i>10.00</i>	<i>10.00</i>	<i>10.00</i>	<i>4.12</i>	<i>3.19</i>	<i>10.00</i>	<i>5.57</i>
Skew	221	41.5	115	37.6	24.5	27	174	59.9	55.3
Kurtosis	4.92e4	3090	1.73e4	2640	4670	4870	3.19e4	7420	6820
Min	-0.996	-0.996	-2.23e5	-133	-232	-210	-733	-186	-158
Max	6.05e5	381	1.26e6	549	378	355	2.54e22	503	477
Range	6.05e5	382	1.48e6	682	610	565	2.54e22	689	635
1 st Q	0.161	0.141	0.744	0.425	0.0189	0.00353	0.00304	0.0507	0.0127
3 rd Q	0.63	0.535	6.82	1.79	0.299	0.115	0.089	0.425	0.159
SD	2720	3.69	7630	6.07	3.5	3.09	1.32e20	4.12	3.79
IQR	0.469	0.394	6.08	1.36	0.28	0.111	0.086	0.374	0.146
MAD	0.295	0.255	2.75	0.828	0.189	0.071	0.0563	0.238	0.0868
CV	173	5.05	60.4	3.31	15	20.1	154	9.71	14.6

Notes

- (1) 10% TM: 10 per cent trimmed mean; 1st Q: first quartile; 3rd Q: third quartile; SD: standard deviation; IQR: interquartile range; MAD: median absolute deviation; CV: co-efficient of variation; GRP: general rate of profit.
- (2) GRP is $s_i / |k_i|$
- (3) Mode.nrd: bandwidth computed by the function `bandwidth.nrd()` in *S-Plus 2000*. Mode.alt: bandwidth as computed by `bandwidth.nrd()`, but capped at 10

Table 3.4 (b): summary statistics for underlying data, accounting ratios

	ORE	ROCE	ATO	NPM
Missing	3.26e4	3.88e4	4.84e4	4.39e4
GRP	0.1756	0.0387	0.8903	0.0700
Mean	286	1.22	2.47	0.678
10% TM	0.248	0.0706	1.84	0.0633
Median	0.174	0.0573	1.74	0.0379
Mode.nrd	0.0514	0.0156	0.1233	0.0112
			(154.14)	
<i>bandwidth</i>	<i>14.50</i>	<i>4.15</i>	<i>62.70</i>	<i>3.86</i>
Mode.alt	0.0358	0.0156	0.0832	0.0112
			(154.14)	
<i>bandwidth</i>	<i>10.00</i>	<i>4.15</i>	<i>10.00</i>	<i>3.86</i>
Skew	147	262	159	110
Kurtosis	2.47e4	6.9e4	2.69e4	1.8e4
Min	-1.47e5	-84.1	-0.0255	-4160
Max	4.12e6	7.08e4	5700	9600
Range	4.27e6	7.09e4	5700	1.38e4
1 st Q	0.0284	0.00999	0.937	0.00514
3 rd Q	0.462	0.132	2.72	0.117
SD	2.25e4	269	31.1	56.2
IQR	0.434	0.122	1.78	0.112
MAD	0.261	0.0818	1.31	0.0633
CV	78.9	221	12.6	82.8

Note

For ATO the secondary mode is given in brackets.

Table 3.4 (c): summary statistics for underlying data, Glick profit rate measures

	glick.1	glick.2	glick.3	glick.4	glick.5	glick.6	glick.7	glick.8
Missing	5.65e4	7.26e4	3.88e4	6.69e4	4.64e4	3.25e4	4.39e4	4.29e4
GRP	0.0429	0.1096	0.0294	0.0794	0.0517	0.1339	0.0500	0.3364
Mean	0.084	-3.92e17	1.18	-3.38e17	0.112	252	0.592	176
10% TM	0.0611	0.168	0.0521	0.132	0.0873	0.182	0.0467	0.684
Median	0.0514	0.125	0.0428	0.0989	0.0783	0.128	0.0279	0.305
Mode.nrd	0.0174	0.0499	0.0118	0.0415	0.054	0.041	0.0082	0.1144
Bandwidth	3.96	11.55	3.12	8.78	3.98	10.75	2.87	37.09
Mode.alt	0.0174	0.0458	0.0118	0.0415	0.054	0.0388	0.0082	0.0884
Bandwidth	3.96	10.00	3.12	8.78	3.98	10.00	2.87	10.00
Skew	212	-148	263	-160	24	160	93.7	88.3
Kurtosis	4.72e4	2.3e4	6.91e4	2.67e4	6470	2.85e4	1.53e4	1.2e4
Min	-84.1	-9.01e21	-84.6	-9.01e21	-84.4	-1.6e5	-4160	-4.56e5
Max	786	6190	7.1e4	3840	85.6	4.12e6	7800	1.51e6
Range	870	9.01e21	7.11e4	9.01e21	170	4.28e6	1.2e4	1.97e6
1st Q	0.00767	0.02	0.00677	0.0159	0.0303	0.0191	0.00332	0.081
3rd Q	0.117	0.317	0.0984	0.248	0.145	0.341	0.0864	1.16
SD	3.54	5.44e19	270	5.05e19	0.773	2.17e4	48	9390
IQR	0.109	0.297	0.0916	0.232	0.115	0.322	0.0831	1.08
MAD	0.0753	0.193	0.062	0.152	0.0812	0.197	0.0481	0.458
CV	42.1	-139	229	-150	6.92	86	81.2	53.3

In the second row we show the Marxian general rate of profit (GRP) – the total net income of the corporate sector divided by the total capital it employs – calculated, in accordance with our discussion above of the treatment of negative capital measures, as $s/|k|$, where s, k , are the net income and capital respectively of each firm, as defined for the profit rate measure in question. We thus calculate the size-weighted mean profit rate and allow comparison of this with unweighted measures of central tendency of company profit rates.

Next come estimates of central tendency: the mean, the ten per cent trimmed mean (found by ignoring the smallest and largest five per cent of observations and calculating the

mean of the remainder), the median, and two alternative kernel density estimates of the mode.³³

The first estimate (mode.nrd) is found as follows. We use the normal reference density (NRD) method – a rule of thumb presented in Silverman (1986: 45-47) – to select the bandwidth of a Gaussian kernel density estimator. However, Gaussian kernels have infinite support and thus the resulting density estimate is necessarily non-zero at every point. This tends to redistribute the probability mass away from the mode and towards the tails, thus reducing the density estimate at the former point. While in principle this should not matter, given appropriate choice of bandwidth (Härdle 1991:78), given the extreme ranges of our profit rate measures we prefer to use a triangular kernel, which is in principle capable of giving a density estimate of zero, and to multiply the NRD bandwidth by 2.432, the transformation factor given by Härdle (1991:76). For ATO we give estimates of both the main and secondary mode. NRD bandwidths are given in *italic*.

However, for certain profit rate measures – in particular Gillman 3 and ATO (the primary mode especially) – the NRD method gives estimates which seem clearly wrong, when compared with the histogram estimates. We therefore give alternative estimates (mode.alt) obtained by estimating the bandwidth with the NRD method but capping the NRD at 10 before transformation. It will be seen that the resulting estimates are considerably more plausible in the light of the histograms.

As inspection of the histograms above would lead one to expect, the estimates of central tendency change sharply as one progresses from the mean to the mode – very sharply indeed, in the cases of Gillman 1, 3 and 6, ORE, and Glick 2 to 4, 6 and 8.

³³ We could estimate the mode simply by taking it to be the bin containing the greatest number of observations, but as examination of Figures 3.3 to 3.6 suggests this leaves one vulnerable to noise in the data, which is not likely to produce monotonic changes in density estimate on either side of the highest bin.

The next pair of statistics are the skewness and kurtosis, both of which have very high values for all measures, in consequence of the extreme ranges of the data and the very lengthy tails to the main probability masses.

The latter point is indicated by the next two groups of data, which indicate first the range and then the dispersion. The nature of the tails is particularly strongly brought out if the standard deviation and coefficient of variation are compared with more robust measures of dispersion, the inter-quartile range and median average dispersion.

A particularly problematic feature is that all these distributions have very extended *lower* tails (with the exceptions only of Gillman 1 and 2, and ATO, though it should be noted that the first two measures mentioned still have minima whose absolute values are quite large in relation to the corresponding general rates of profit; bear in mind also the exceptional status of ATO on account of its bi-modality).

This is important because the two distributions proposed as models for the rate of profit – the log normal and the gamma – both have bounded support (bounded below at zero, in the case of their standard forms). Extended left-hand tails are either inconsistent with such models, or imply parameter values such that the distributions are approximately normal in any case. We will return to this topic in Chapter Four.

3.3 Extreme values and firm size

Further insight is given by considering the extreme tail quantiles for each profit rate distribution: the minimum value (zero quantile), the 0.5, one, two, five and ten per cent quantiles, and the symmetrical quantiles from the upper tails. These are detailed in Table 3.5 (10 and 90 per cent quantiles omitted for space reasons), but their importance is more clearly seen from graphing the inter-quantile ranges, as is done in Figures 3.7 to 3.12.

Table 3.5: tail quantiles of firm-level distributions

PRM	0.0%	0.5%	1.0%	2.0%	5.0%	95.0%	98.0%	99.0%	99.5%	100.0%
p.gill.1	-0.996	-0.176	-0.0626	0	0.0392	3.07	9.08	2.03e1	4.81e1	6.05e5
p.gill.2	-0.996	-0.281	-0.145	-0.0383	0.0312	2.02	4.68	8.92	1.49e1	3.81e2
p.gill.3	-2.23e5	-1.66	-0.372	-0.0573	0.0769	3.38e1	8.24e1	1.68e2	3.68e2	1.26e6
p.gill.4	-1.33e2	-0.346	-0.121	-4.14e-3	0.0851	6	1.06e1	1.62e1	2.39e1	5.49e2
p.gill.5s	-2.32e2	-3.78	-2.1	-1.11	-0.406	1.21	2.34	3.88	6.34	3.78e2
p.gill.5f	-2.1e2	-1.97	-0.96	-0.473	-0.167	0.54	1.44	3.29	6.59	3.55e2
p.gill.6	-7.33e2	-0.664	-0.454	-0.276	-0.113	0.351	0.782	1.6	2.85	2.54e22
p.gill.7s	-1.86e2	-3.02	-1.7	-0.861	-0.289	1.67	3.26	5.39	8.69	5.03e2
p.gill.7f	-1.58e2	-1.31	-0.702	-0.355	-0.117	0.731	1.98	4.61	8.84	4.77e2
p.ORE	-1.47e5	-9.62	-4.26	-1.94	-0.684	2.77	1.4e1	5.58e1	2.33e2	4.12e6
p.ROCE	-8.41e1	-0.799	-0.509	-0.305	-0.127	0.383	0.752	1.36	2.4	7.08e4
p.ATO	-0.0255	8.96e-3	0.0229	0.0445	0.0884	5.55	8.17	1.12e1	1.59e1	5.7e3
p.NPM	-4.16e3	-2.04	-0.828	-0.384	-0.119	0.713	1.07	2.32	5.83	9.6e3
p.glick.1	-8.41e1	-0.803	-0.52	-0.32	-0.137	0.303	0.477	0.753	1.23	7.86e2
p.glick.2	-9.01e21	-5.71	-2.88	-1.41	-0.496	1.41	3.22	6.26	1.22e1	6.19e3
p.glick.3	-8.46e1	-0.768	-0.474	-0.278	-0.115	0.295	0.599	1.09	1.88	7.1e4
p.glick.4	-9.01e21	-4.73	-2.46	-1.22	-0.419	1.13	2.69	5.46	1.17e1	3.84e3
p.glick.5	-8.44e1	-0.697	-0.424	-0.236	-0.0827	0.333	0.592	1.09	1.8	8.56e1
p.glick.6	-1.6e5	-8.7	-3.8	-1.76	-0.634	2.09	1.06e1	4.22e1	1.81e2	4.12e6
p.glick.7	-4.16e3	-1.98	-0.816	-0.364	-0.11	0.551	0.974	2.15	5.87	7.8e3
p.glick.8	-4.56e5	-2.42e1	-9.75	-3.81	-0.811	1.3e1	4.9e1	1.37e2	3.97e2	1.51e6

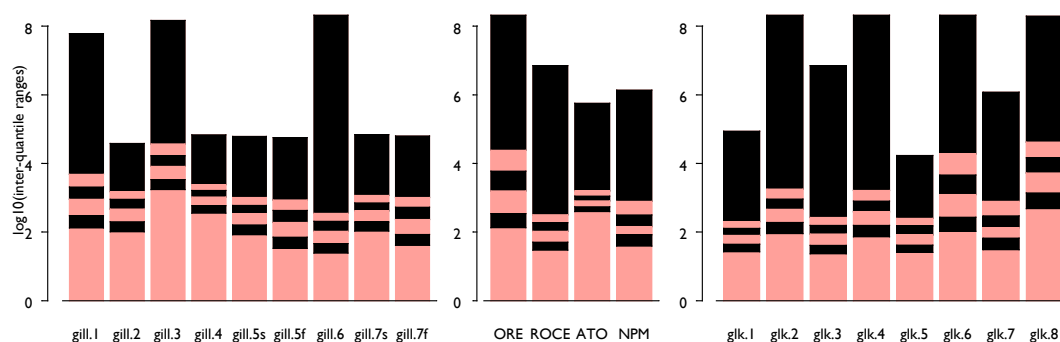
**Figure 3.7: log ranges of profit rate measures; extreme bars truncated**

Figure 3.7 compares the *logarithms* (to base 10) of the inter-quantile ranges reported in Figure 3.5 for each profit rate measure. The total height of each bar represents the log of the range of the whole data; the ranges resulting from progressive trimming of the data are represented by discarding successive sub-divisions: thus, the range of the middle 99 per

cent of the data can be seen by the height of the bar less the topmost sub-division (solid colour), the range of the middle 98 per cent by the height less the top two sub-divisions (solid and shaded), and so on. To better appreciate the behaviour of the distributions in the regions where the tails join the main mass we truncate the log range. As can be seen, although the total ranges are great, the inter-quantile ranges decrease sharply in most cases: trimming the outer ten per cent of observations reduces the range very sharply in all cases. Once again, note that because we graph the *log* scales, the length of the tails is in fact far more extreme even than indicated here. Among other points, note that it is now clear that the central 80 per cent of observations of Gillman 6 have a *smaller* range than does the same proportion of observations of most other measures, notwithstanding its freakish overall range.

We now show that the long tails seen for each profit rate measure are due to the existence of substantial size effects: the range of any given profit rate measure associated with small firms is larger than the range associated with larger ones.

To do this we group firms by order of magnitude of the capital measure used in each rate of profit measure, and plot the tail quantiles of Figures 3.7 and 3.8 for each size class and for each profit measure. Thus Figures 3.8 to 3.11 show (on \log_{10} scales) the overall range, together with the central 99, 98, 96, 90 and 80 per cent of the data, stratified by order of magnitude of the relevant capital measure.

Since our subject is capital and its (maximum) rate of expansion, the rate of profit, our method of classifying firms by the amount of their capital seems to us both obvious and natural. Nonetheless, it may appear controversial to some: Serrano Cinca *et al.* (2001) point out that the European Commission's recommendation 96/280/CE gives priority to number of employees in the definition of size, while they themselves claim that there is 'general agreement' that for the purposes of economic research the effect of size is best captured by turnover (Serrano Cinca, 2001:2, and see further references there).

Since the purpose of their paper is precisely to identify size effects in financial ratios the difference between their choice of classification system and ours is potentially far from trivial. We will discuss their results when reviewing our own.

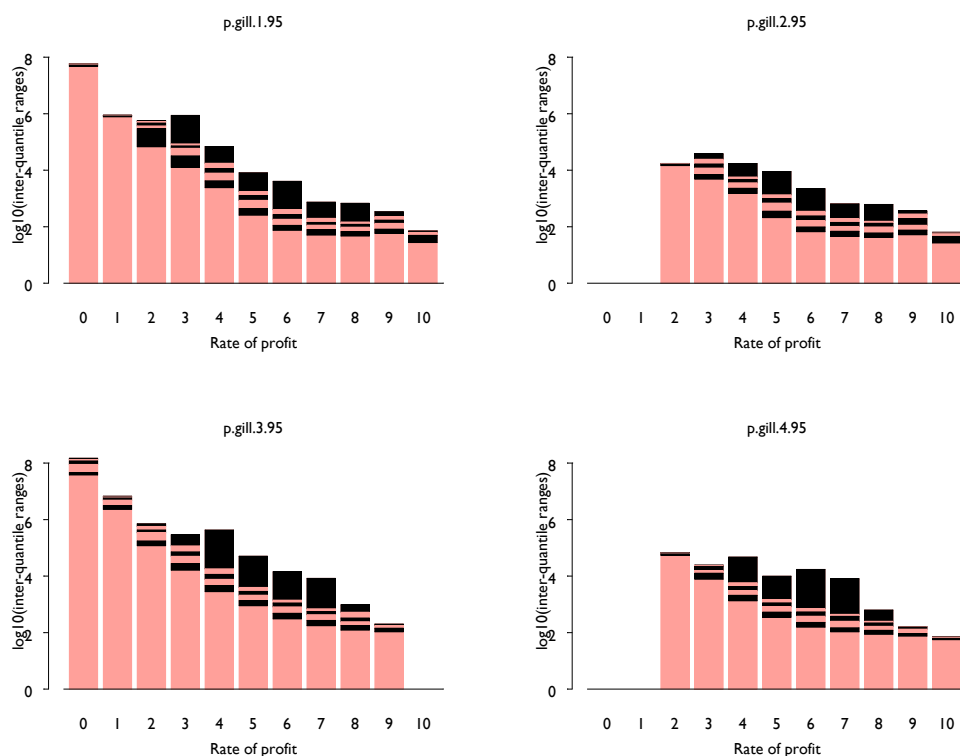


Figure 3.8: firm size and range of profit rate, Gillman's 'marxian' measures

We first consider Gillman's 'marxian' measures (Figure 3.9). As can be seen, for each measure the range of profit rates falls steadily as one moves up from the smallest magnitude of firm size. There is also some tendency for the middle ranges of firm size still to exhibit extended tails, but they are clearly much less so than the tails for the whole population of firms. Note that the quantiles plotted are those of the size category in question, not the quantiles of the whole population; thus, in the case of Gillman 1, for example, the range of the central 80 per cent of firms with capital between £0 and £10 is greater than the range for all companies with capital in the £11-£20 range.

This inverse relation between profit rate range and firm size is a feature of all measures, as Figures 3.9 to 3.11 show.

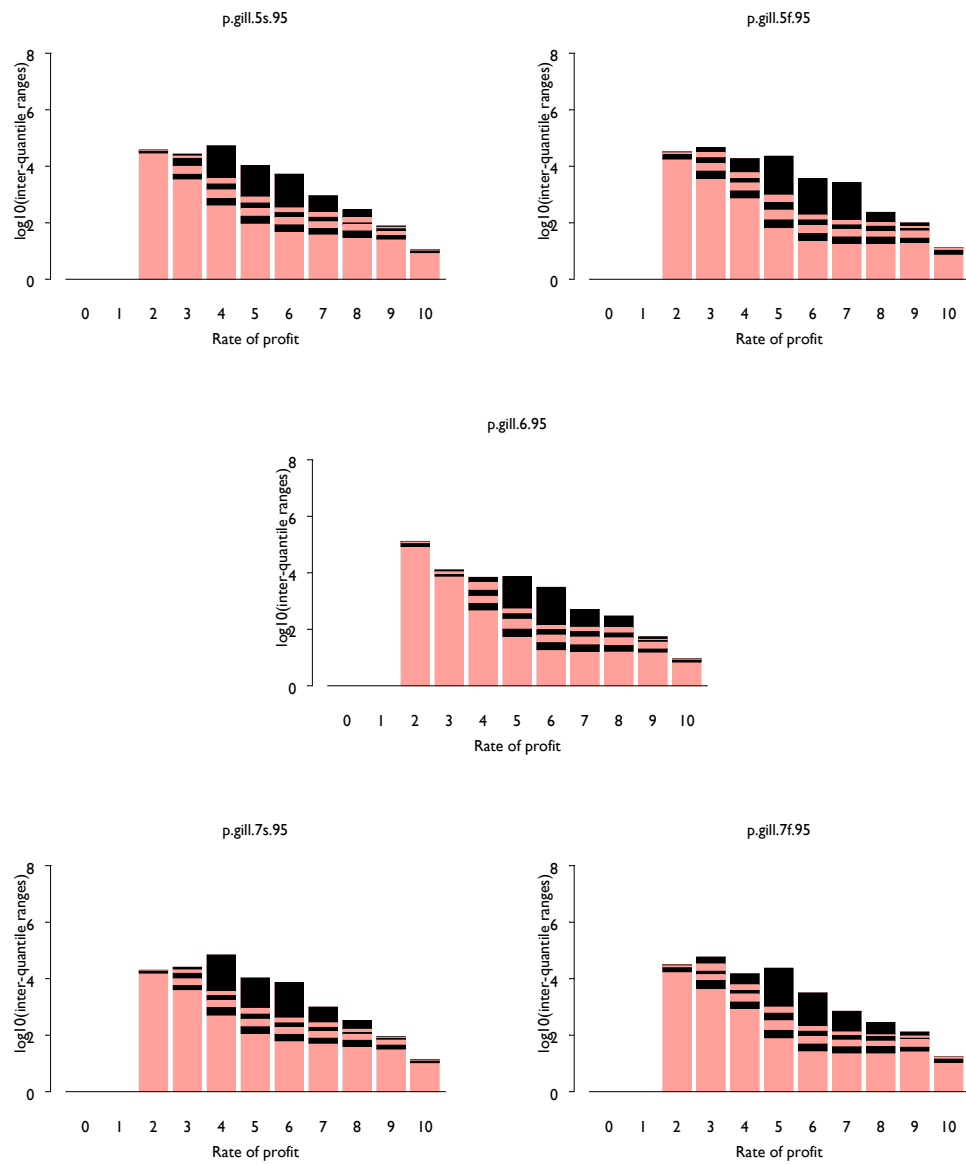


Figure 3.9: firm size and range of profit rate, Gillman's 'capitalist' measures

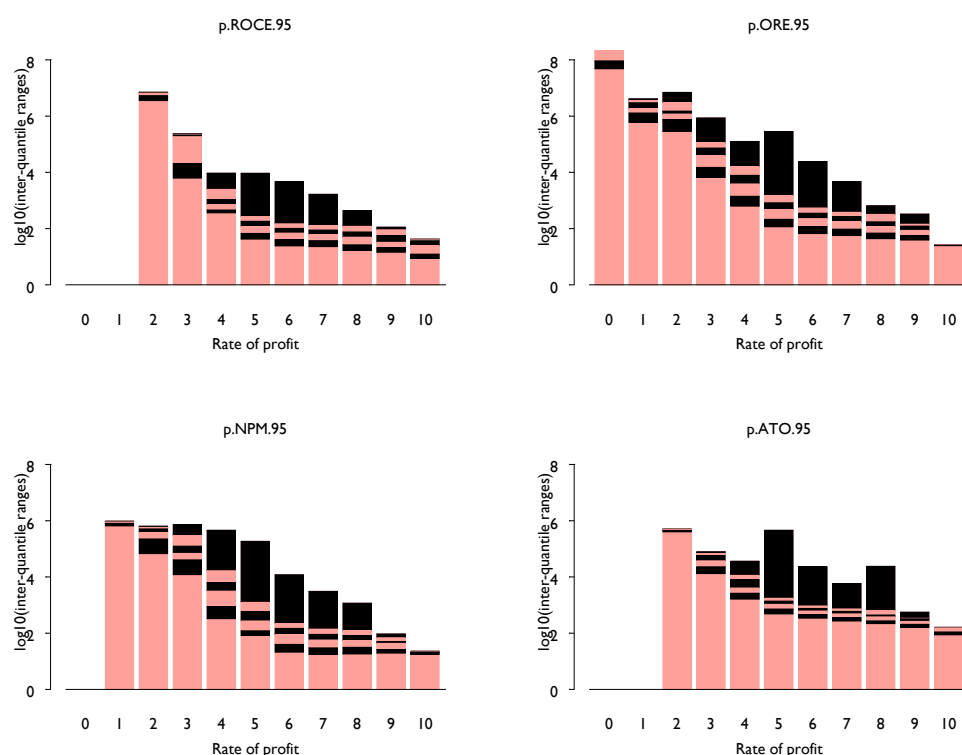


Figure 3.10: firm size and range of profit rate, accounting ratios

We note particularly that the size effect on range of profit rates for the four accounting ratios (Figure 3.10) is at least as strong, if not stronger, as that for other profit rate measures.

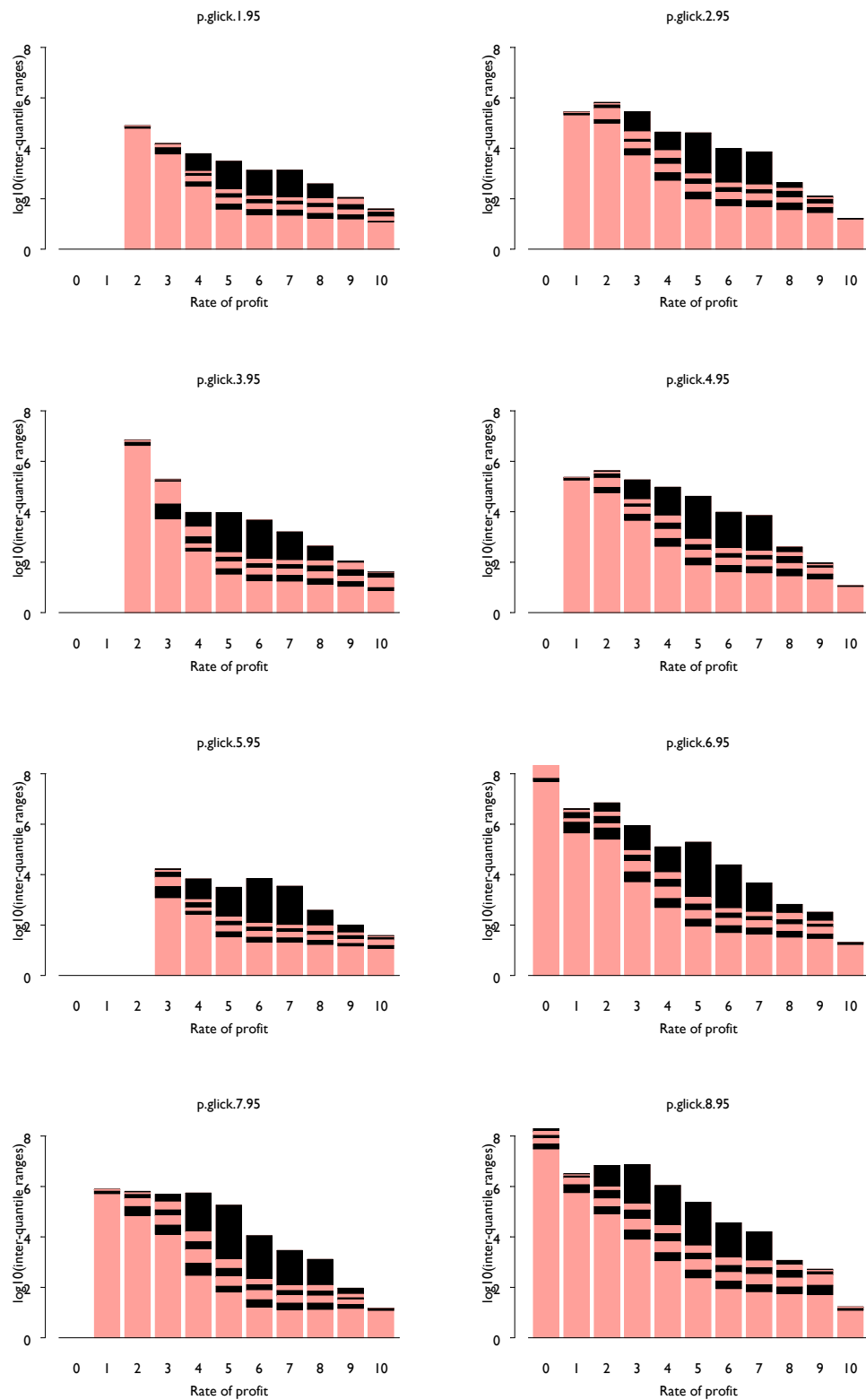


Figure 3.1 I: firm size and range of profit rate, Glick measures

As Figures 3.8 to 3.11 show, we find size effects to be common to all profit rate measures, including the four accounting ratios, and ORE in particular. These results are at variance with those of Serrano Cinca *et al.* (2001). They found that the return on equity (our ORE) had no power to discriminate between groups of firms of different sizes.³⁴

In contrast, both gross and net profit ratios (the latter equivalent to our NPM) did have such power, along with other ratios such as staff costs relative to value added (akin to the rate of surplus value) and the ratio of goods and services purchased to value added (which is connected to the organic composition of capital). Their conclusion was that these results indicated different choices of strategy by differently-sized firms aimed at the same end, namely financial profitability as measured by ORE (Serrano Cinca 2001:12). In the language of the present study this is close to saying that inter-firm differences in organic composition of capital (and in rate of surplus value, if they exist) nonetheless lead to an approximately equalised rate of profit (as measured by ORE)

Their methodology is very different to ours; they use normal ANOVA techniques, as well as discriminant analysis and cluster analysis, and work with industry level data. And as noted above, their classification of size is on the basis of turnover (with data on the proportion of firms in each industry falling into each of three groups: turnover less than seven million Euros, even to 40 million, and greater than 40 million Euros).

Nonetheless, in spite of all differences their finding suggests the following, in reflecting on the empirical density functions of Figures 3.4 to 3.6: the two measures which we take to best represent the rate of profit which is the subject of the transformation problem are Gillman 3 and 4. These measures have distributions with forms which are, subjectively, clearly distinct from those of all other measures (bar the idiosyncratic ATO). If these other measures, including as they do information about firms' financial structure and the

³⁴ For example, they found the mean of this ratio to be virtually identical across size groups, using data from 15 countries ((Serrano Cinca 2001:23, Table 1).

capitalists' 'other grounds for compensation', as discussed by Marx, then we have a counterpart to the Serrano Cinca result.

In connection with this we note that Marx's value-theoretic account of 'capitalist communism' (the redistribution of value from companies with low organic composition to ones with large OCC) provides an explanation for the smaller range of profit rates attained by large companies. The maximum rate large companies can achieve is constrained by the finite amount of surplus available for redistribution: suppose a total surplus of 100, to be distributed among nine companies with a capital of ten, and one company with a capital of 100: evidently the large firm cannot achieve a profit rate of more than 100 per cent, even if it manages to appropriate the whole surplus, whereas if one of the small companies was to do this its profit rate would be 1,000 per cent.

3.4 Conclusion

Our exploratory investigation of the profit rate measures constructed from FAME's company accounts data has achieved several things, some of primarily technical significance, but others which constitute a substantive contribution in their own right

On the technical side, we have demonstrated that direct use of large company accounts data sets to investigate profit rate distributions at the company level will have to deal with data containing many extreme values.

Substantively, we have provided what to our knowledge is the first demonstration that different types of profit rate measure have different empirical density functions. We have also seen that our work here may be linked to other work on size effects in financial ratios.

We have at the same time demonstrated in cross-section a phenomenon accepted as a stylised fact about time-series data, namely the greater variance in the rates of return of small as opposed to large companies.

However, the above is only an exploratory analysis; our full agenda requires investigation of what, if any, functional form or forms govern the empirical distributions shown above. The extended tails of these distributions make identification and estimation of such models harder. Methods aimed at addressing this issue will be developed in the next chapter.

Chapter 4 Methods

This chapter develops a methodology for empirical investigation of hypotheses about the distribution of profit-rate measures. In principle we would like to test hypotheses about distributions of the profit rate both across firms (for consistency with Gibrat) and across units of capital (for consistency with Farjoun and Machover). A possible method for estimating firm-level densities, given the nature of the data, reduces to a method for estimating capital-level densities.

In the previous chapter we saw that our profit rate measures exhibited long dense tails. This chapter considers how this long tail issue might be addressed. In Section 1 we discuss the need to distinguish whether observations are outliers, extreme values, contaminants, or discordant (or some combination of these), and thus the need for different strategies to deal with these different issues. In Section 2 we discuss a possible estimation method, the method of L -moments, that is not only robust to the presence of extreme values but offers a convenient answer to a particular problem of discordancy in our data.

In addition, Section 3 argues that much that is problematic about the data is due to it being derived from two different populations, only one of which is of interest. Weighting the data by firm size is an appropriate, non-arbitrary method of selecting observations for exclusion, while also providing accommodation of genuine extreme values. It is also provides an estimate of the distribution of profit rates across the capital employed, rather than across firms.

In Section 4 we begin by calculating L -moments directly from the firm level data on our 21 profit rate measures. Section 5 uses size-weighted random samples to estimate L -moments and thus identify possible distributional models for the profit rate measures at the capital level. Finally, Section 6 concludes that this procedure is also an appropriate method for estimates which accommodate the extreme features of the firm-level data.

4.1 Extreme values, outliers, and discordancy

4.1.1 Review of issues

At this point it is helpful to recall some distinctions drawn in the standard text on outliers (Barnett and Lewis, 1994). First, we note their definition of an outlier: ‘an observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data’ (page 7). The authors stress that the phrase ‘appears to be inconsistent’ is crucial; what worries the researcher is whether the observations are genuine members of the main population, or whether they are *contaminants* – members of some other population which have somehow crept into the sample. Fundamentally, the unreasonableness of a suspect observation is entirely relative to some model of the distribution.

The problem of outliers is often associated with extreme values which appear to be unreasonably distant from the main body of observations. Yet, as Barnett and Lewis show, ‘outlier’, ‘extreme value’ and ‘contaminant’ are three distinct concepts, and it is perfectly possible (as their Figure 1.3, page 9, shows) for an observation to be both a contaminant and an extreme without being an outlier, or to be an extreme outlier without being a contaminant.

Extreme values are the maximum (or minimum) of some data; but they will not be outliers if, for example, they are drawn from a standard Gaussian distribution and fall within one standard deviation of the mean.

Outliers are observations which are (surprisingly) distant from the main data, but if there are many such, plainly only one (in a given tail) can be an extreme. In panel (a) of Figure 4.1 the ordered sample statistics $x_{(1)}$ and $x_{(n)}$ are extremes, but if it is thought that they are drawn from distribution F they will not be regarded as outliers. On the other hand in panel (b) $x_{(n)}$ is an outlier because it falls in a region where F has very low density, while $x_{(1)}$ may also give cause for concern.

Contaminants come from other distributions. They may (or may not) be extreme (or outliers) either in relation to their own distribution, or the ‘host’ data. In panel (c) of Figure

4.1 there are two such contaminants, denoted (\bullet) from the distribution G , of which one is in the middle of the sample and the other is the upper extreme. But although $x_{(n)}$ is both a contaminant and an extreme it is not an outlier. In contrast, in panel (d) the contaminant is non-extreme, but is a member of the outlier pair $x_{(n-1)}, x_{(n)}$.

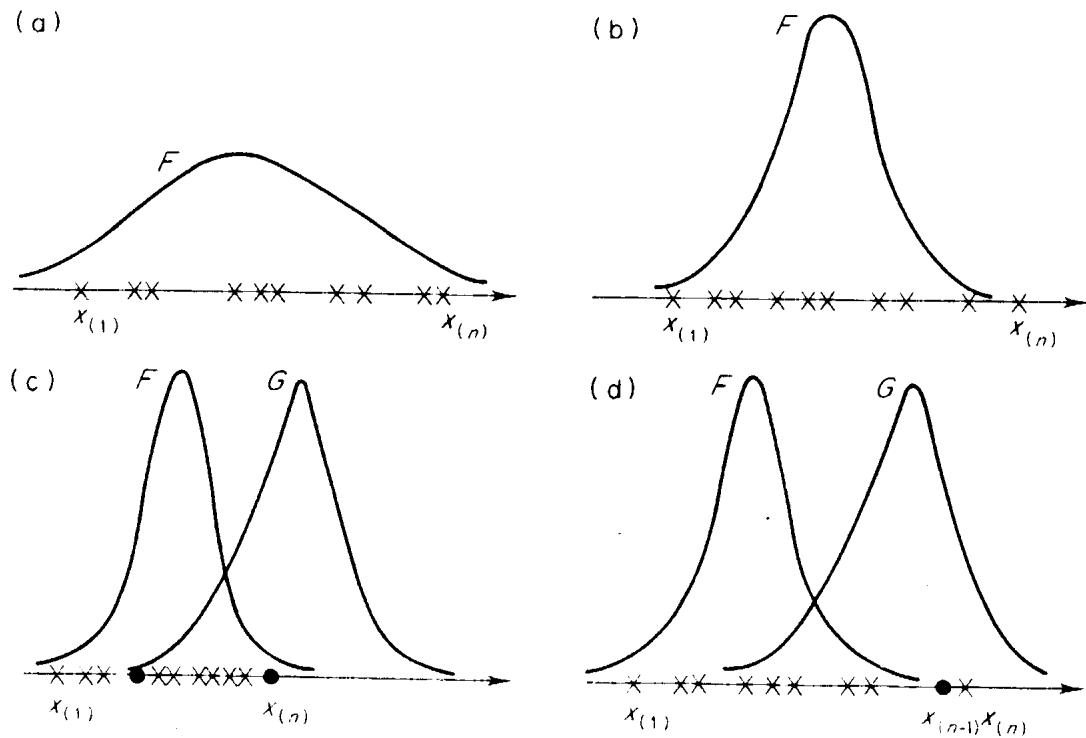


Figure 4.1: extremes, outliers and contaminants (Barnett and Lewis, 1994: 9)

Barnett and Lewis suggest four possible responses to the presence of outliers (page 35):

- (1) *accommodation*, in the sense of using analytical tools that are robust to the presence of outliers
- (2) *rejection* and thus elimination of the offending data
- (3) *incorporation* in a modified model
- (4) *identification* of 'hitherto unsuspected factors of practical importance'.

The latter two may appear similar, but the distinction Barnett and Lewis appear to have in mind is between *incorporation* in a better descriptive model – in other words, a particular distribution function that is a better fit to the data – and *identification* of a new causal

model, in other words, a revised view about whatever causes the data to have the particular distribution function which describes it.

Which response is adopted will depend on what we believe about how the outliers have arisen. Again, four main categories may be distinguished (pages 45ff).

(a) *Deterministic* causes are ‘obvious gross errors of measurement, recording, and so on’. In such cases the appropriate response is clearly (1) *rejection* of the suspect data.

(b) *Inherent variability* attributes the presence of outliers to the fact that the variability of the population simply has different characteristics than supposed by our original model; in this case, we will (3) *incorporate* the outliers in a modified model, for example with fatter tails or skewness in place of symmetry.

(c) *Contamination* is a situation in which the outliers are observations drawn from another population with a different distribution. We may be dealing with a *mixture* of distributions, where that describing the contaminants is either of different form entirely or from a distribution of the same kind but with different parameters (so-called slippage models). Here one faces the possibility that further, non-outlier observations are also contaminants. In principle any of (2) rejection, (3) incorporation, or (4) identification of new factors might be appropriate.

(d) *Proneness* of the given distribution to outliers. The *absolute* and *relative* proneness of distributions to exhibit outlying observations are defined by the probabilities of exceeding the difference and the ratio, respectively, of the two most extreme values (Green, 1976).

In the context of testing Farjoun and Machover we note that the gamma distribution, although *absolutely* outlier prone in the sense just mentioned, is *relatively* outlier resistant (as are log normal distributions; Gaussian distributions are resistant in both senses; Cauchy distributions – found, for example, as the ratio of two Gaussian distributions – are an

example of proneness in both senses). In this case the appropriate strategy must be (1) accommodation.

One possible problem is that we may face more than one source of outliers, in which case several strategies may have to be combined.³⁵

4.1.2 Sources of outliers in profit rate data

The summary statistics presented in Chapter Three suggest not only that each of our measures contains very extreme values, but that these may be outliers in the sense of being distant from the main body of the data. Of perhaps greater importance is the fact that large subsets of each data set (and not merely the two extreme values in each) are outliers in the Barnett and Lewis sense of being surprising in relation to particular models.

In this section we consider which of the sources of outliers described above may be relevant to our data.

4.1.2.1 Deterministic causes

One possible source of data error, which if present could be very large, arises from the way FAME data is presented. Normally the data is rounded to the nearest £1,000 and the trailing zeros omitted; but for firms where no figure exceeds £10,000,000 the data is expressed in actual figures (the form adopted for each firm is identified by an indicator variable UNIT).

Clearly any inconsistency in doing this for different variables within a particular record could lead to errors of the magnitude of 10^3 (either up or down, depending on whether the direction of the error and whether it occurs in an item appearing in the numerator or denominator of a particular profit measure). In practice, any such errors might be of lesser

³⁵ However, apart from pointing out the link between deterministic causation of outliers and rejection, Barnett and Lewis do not explicitly discuss how causes might correspond to responses.

magnitude, in that some measures involve numerators and/or denominators calculated from two or more variables; in these cases, the order of magnitude would be smaller, unless one supposes that the errors occur in exactly the variables forming the numerator or denominator, as the case may be.

Since any such errors may be as likely to be downward as upward, they could thus lead to discordantly small profit rate observations as well as ones which are discordantly large. The first, however, will be hard to detect, being mixed in with genuine observations of very small profit rates.

However, while the above establishes the *possibility* of very large recording errors, it says nothing about the *probability* of their occurring. We have no reason, apart from the above arguments, to suspect their presence. Since the FAME database is prepared by a commercial entity and primarily used by commercial customers who will value accuracy highly (because they may, for example, base lending and other investment decisions on the results) it seems that persistent presence of many errors is unlikely.

4.1.2.2 Inherent variability

The possibility of the outliers being due to some model of variability other than those of our two hypotheses is inherent in the notion of testing them.

4.1.2.3 Contamination

As a large literature descending from Gibrat attests (Sutton 1997), the distribution of firms by size is highly skewed by almost any measure of size. The same is true if we use the definitions implied by the various profit-rate measures examined in this paper.³⁶

³⁶ At this point, recall our discussion of Serrano Cinca *et al.* (2001) in Chapter Three.

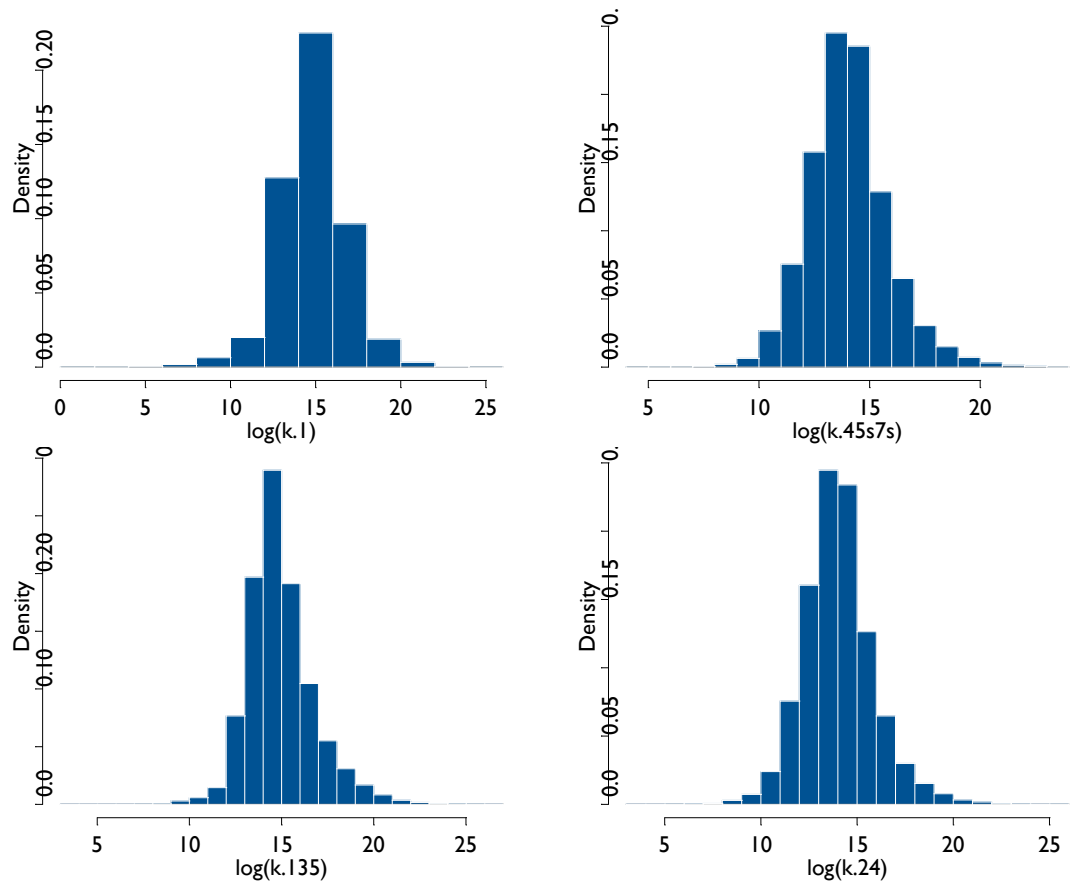


Figure 4.2: firm size distributions, selected capital measures (see text)

Figure 4.2 present histograms of the *log* of firm size, measured by four different capital measures (from left to right in each row, those involved in the profit rate measures Gillman 1; Gillman 4, 5s and 7s; Glick 1, 3 and 5; Glick 2 and 4). Even in these logarithmic plots some skewness is apparent, with each plot showing an upper tail longer than the lower.

Moreover, whatever measure of ‘size’ is used, the range is vast – for example, there are firms whose cost of sales is measured in mere thousands of pounds alongside ones with cost of sales in the hundreds or thousands of millions, in other words bigger by a factor of 10^6 or more.

Many such small firms are not genuine capitalist enterprises, but tax vehicles for individual contractors or enterprises which employ only family members.³⁷ If homes or cars can be put in the firm's name, for instance, its net assets could easily be of the order of £10⁵. As such, the profit and capital figures in the accounts may be driven by tax considerations rather than a need to accurately measure rates of return on capital and we should not be surprised at very extreme values (perhaps especially not at highly negative ones, if individuals have incentives to under-report both income and assets). We will want to exclude these observations as contaminants.

However, there is also the question of entities which, although small, are genuinely capitalistic. In Chapter Three Section 3 we saw that the size of firms has a strong inverse relation to the range of profit rates which they experience. One possible explanation is that this is a case of mixture of distributions, with slippage of the scale parameter, and possibly others. But pursuing this line meets with two objections: first, we do not, in the nature of the case, have any information about what distributions might be mixed; second, if at all possible we would prefer a simpler model, in the sense of a single distributional law covering all cases. We will take up this theme again in Chapter Seven.

4.1.2.4 Proneness to outliers

One possible source of outliers is that the theoretical model is intrinsically prone to them. This is in fact the case with the gamma distribution (absolutely, but not relatively, as already noted).

However, even if appropriate tests showed that the right tail observations were not discordant with this hypothesis, we still have to account for the left tail: since the gamma and log normal distributions have bounded support, some part of this *must* be discordant,

³⁷ In 1995, of the 54,285 enterprises which reported the number of employees, 433 reported just one; 1,720 reported two or fewer; 5,198 five or fewer; and 9,599 reported 10 or fewer

unless we fix the threshold at or to the left of the leftmost observation – or unless we adopt some third model.

As noted in the introduction to this chapter, we will be adopting methods which are capable of automatically suggesting alternative candidates.

4.1.3 Strategies for outliers in profit rate data

In this section we consider strategies for each of the sources of outliers described above.

4.1.3.1 Deterministic causes

Unfortunately we have no way of identifying which, if any, data may be subject to such errors. So we discount the possibility of simply rejecting observations on the grounds of unlikelihood and impracticability, and rely on methods robust to the presence of outliers.

4.1.3.2 Inherent variability

As will be seen in Section 4.2.6, the methods that we will adopt will automatically indicate whether an alternative should be considered (though not necessarily what that alternative should be).

One possibility is the skewed- t distribution introduced by Jones and Faddy (2003), in which the weight of each tail is controlled by a distinct parameter; this could be fitted by maximum-likelihood methods. But our only motive for considering this is the possibility of descriptive accuracy; we do not have any theoretical reason to suppose it valid.³⁸

Other possibilities include special cases of generalised versions of the extreme value, logistic, and lognormal distributions. However, even if we overlook the absence of

³⁸ Moreover, the methods we will be using cannot readily identify and estimate the skewed- t .

theoretical motivation, their possibilities for descriptive accuracy are limited by their fixed shape.³⁹

A further alternative, attractive for its ability to model distributions with extreme skewness and kurtosis, is the stable family of distributions. This will be discussed further in Chapter Seven.

4.1.3.3 Contamination

There are two problems under this heading: first the possibility that some of our data relates to non-capitalist entities with essentially arbitrary rates of return, and second the presence of genuinely capitalist enterprises whose small size is associated with one or more distributions different to those of larger firms.

In either case the firms in question may record profit rates with small absolute magnitudes as well as extreme values, so we need a criterion by which to identify the relevant observations and either exclude or accommodate them even though they may be neither outliers nor extreme values.

There are various approaches one might take. First, one might exclude all firms with only one employee, in the hope that that will dispose of firms that are simply vehicles for independent contractors. Secondly, as we saw in our discussion of Serrano Cinco *et al.* (2001) in Chapter Three, Section 3, there are various *ad hoc* definitions of small and medium sized enterprises, some conventional in the literature and others devised for various official purposes, including the collection and classification of data.

One might either use one of these as further grounds for excluding observations, or as the basis for a formal mixture-of-distributions model and test for that. Thirdly, one might

³⁹ Moreover, if they do happen to be valid they will automatically be detected by methods to be described below.

weight the data by firm size before testing, which would be tantamount to estimating what we have called the capital-level or capital space distribution.

We would definitely like to exclude non-capitalist entities: theoretically, we do not believe that they participate in capitalist communism; practically, we believe that their reported rates of return are arbitrary constructs. On the other hand, we would like to accommodate small capitalist firms in spite of their greater range of profitability. In either case, we would like to avoid arbitrary criteria as far as possible. If we had some clear theoretical basis for some exclusion criteria this would amount to a discrete, all-or-nothing weighting scheme for dealing with small firms.

But we know that variability in profit rates is an inverse function of firm size (see Chapter Three, Figures 3.9 to 3.12). Observations of small firms' profit rates are likely to be contaminants, in the sense that such entities may not be capitalist firms, in which case we want to reject them. If we take the view that the probability of an observation's being such a contaminant is inversely related to firm size, then a system in which an observation's probability of inclusion in our tests depended directly on the size of the firm to which it relates constitutes a system of fuzzy rejection of (possible) contaminants. In the words of an author cited by Barnett and Lewis (1994):

Since the object of combining observations is to obtain the best possible estimate of the true value of a magnitude, the principle ... [is to assign] a smaller weight than the others in computing a weighted average. Of course retention with an exceedingly small weight amounts to virtual rejection. (Rider, 1933)

In place of 'virtual rejection' we prefer 'fuzzy rejection', since not all the weights are 'exceedingly small', and in fact some will be close to 1 (and one exactly 1). (See also Edgeworth, 1883, Glaisher, 1872–3, Stigler, S.M., , Stone, 1873). Stone is effectively maximum likelihood estimation, and Edgeworth an independent discovery of the same.

4.1.3.4 Proneness to outliers

There are two aspects to this. The first is the tendency of some distributional models to throw up outliers in the right-hand tail, which may thus be striking, but are not discordant. The second is that, because both our candidate distributions have bounded support, observations below some threshold *must* be discordant. Since we do not have a clear theoretical basis on which to select such a threshold it becomes another parameter to be estimated.

As mentioned in our survey of previous contributions, the only known hypotheses about the functional form of profit rate distributions are the lognormal (Gibrat 1931) and the gamma (Farjoun and Machover 1983). Both these distributions have only bounded support on the real line, in other words they have lower thresholds. Although Farjoun and Machover claim that a lower bound of zero is to be expected (on the grounds that firms which make losses will quickly cease to be firms), it is clear from our empirical density estimates that although the proportion of firms making moderate losses is small, it is nonetheless not negligible.

Furthermore, casual observation suggests that many firms can go on making losses for some time before the creditors get restive enough to call in the receiver. In fact, what the limiting rate of loss might is a quantity estimated in estimating the threshold parameter of a distribution that successfully models the rest of the data.

While the empirical density estimates of Figures 3.3 to 3.6 might suggest that this lower bound is only moderately negative, this appearance is clearly belied by the overall data. On the other hand, as Table 3.5 shows, very moderate trimming of the left-hand tails – only half a per cent of the total observations in many cases, and one or two per cent in nearly all the others – would give minima which are much more plausible as a estimate of a rate-of-loss which might be generally sustainable in the medium term.

Thus any attempt to test the Gibrat and Farjoun and Machover hypotheses by fitting the relevant models to the data involves estimating the threshold parameter along with the

shape and scale parameters.⁴⁰ Unfortunately the usual maximum likelihood methods cannot be relied on to do this (see Appendix).

It would possible by *fiat* to take the minimum observed value x_0 as the estimated threshold. But doing so on the basis of the data in Table 3.4, with their dramatically extended left tails, will result in estimated gamma or log-normal distributions closely approximating the normal, which on the face of Figures 3.3 to 3.6 is *not* a very appropriate model given the pervasive qualitative evidence of skewness; furthermore it will clearly fail to capture the kurtosis (since this is fixed at 3 for the normal).

However, a recent development in statistical technique – the method of L -moments – not only provides simple methods for threshold estimation but also offers more reliable methods for model selection than those based on the use of traditional moments.

4.1.3.5 Summary

We would like to model the distribution of profit rates across both firms and capital, in order to test Gibrat and Farjoun and Machover respectively. Above we have discussed two basic strategies for dealing with the problem of extreme values: methods robust to outliers, and weighting the data as a method of fuzzy rejection of contaminants.

In the following section we consider a method which is not only robust in the presence of extreme values but will solve the problem of estimating location parameters which we identified in our discussion of discordancy.

We then turn to our scheme of fuzzy rejection, equivalent to estimating the capital-level distributions, and show that it diminishes the extreme value problem to the extent that our robust method now has some purchase on the problem.

⁴⁰ There are unbounded distributions which exhibit skewness, some of which are special cases of models which will be considered below. An alternative would be Jones's skewed- t , mentioned earlier, but we do not know of any theoretical motivation to consider this distribution.

4.2 Robustness: the method of L -moments

4.2.1 Definition and description

The measures of skewness and kurtosis presented in Chapter Three, Table 3.4, were the product-moment based measures traditionally used to fit distributions in the Pearson system. They are usually defined as follows: the mean and the higher moments

$$\mu_r = E(X^r); \quad \mu_r = E(X - \mu)^r, r = 2, 3 \dots$$

are the basis of the dimensionless measures of skewness and kurtosis

$$\gamma_1 = \mu_3 / \mu_2^{3/2}; \quad \gamma_2 = \mu_4 / \mu_2^2.$$

An alternative system (Hosking, 1990) is the method of L -moments, a development of so-called probability-weighted moments (p-w moments) introduced by Greenwood *et al.* (1979). The following discussion summarises Hosking and Wallis (1997), which largely supersedes Hosking's earlier article.

4.2.2 L -moments and order statistics

Hosking and Wallis (1997) give an intuitive justification of L -moments by considering linear combinations of a sample of data which has been arranged in ascending order ('order statistics').

Let $X_{k:n}$ be the k^{th} smallest observation from a sample of size n , so that the ordered sample is $X_{1:n} \leq X_{2:n} \leq X_{3:n} \dots X_{n:n}$. A sample of size 1 is the single observation $X_{1:1}$, which contains information about the location of the distribution; if the latter shifts towards larger values then we would expect to see larger values of $X_{k:n}$. A sample of size 2 contains two observations $X_{1:2}$ and $X_{2:2}$, and these give information about the scale: if the distribution is tightly bunched round a central value, then we expect the $X_{k:n}$ to be close together, and if not, not; thus the difference $X_{2:2} - X_{1:2}$ is a measure of scale.

Likewise, in a sample of size 3, $X_{1:1}, X_{2:2}, X_{3:3}$; we expect $X_{3:3} - X_{2:3} - X_{2:3} + X_{1:3}$ and thus the central second difference $X_{3:3} - 2X_{2:3} + X_{1:3} = 0$ if the distribution is symmetrical, whereas this quantity will be expected to be positive if the upper tail is heavier than the lower (right-skewed) and negative in the opposite case.

With the central third difference, $X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}$, re-writing this as $(X_{4:4} - X_{1:4}) - 3(X_{3:4} - X_{2:4})$ shows that this is a measure of the degree to which the two extreme values are further apart than the two central values; in the case of a nearly-flat distribution (nearly uniform) the sample values will be approximately equally spaced and the third central difference close to zero, whereas if the distribution has a high peak then $3(X_{3:4} - X_{2:4})$ will be small relative to $X_{4:4} - X_{1:4}$; thus it is a measure of the kurtosis. In this case the L -moments are

$$\begin{aligned} L_1 &= E(X_{1:1}) \\ L_2 &= \frac{1}{2} E(X_{2:2} - X_{1:2}) \\ L_3 &= \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}) \\ L_4 &= \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) \end{aligned}$$

and in general

$$L_r = \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E(X_{r-j:r}).$$

Dimensionless versions are given by dividing the higher L -moments by the scale measure L_2 , so that $L_r = L_r / L_2$; $r = 3, 4, \dots$. Additionally, a quantity analogous to the usual coefficient of variation is given by L_3 / L_1 .

Hosking and Wallis (1997) uses the abbreviation L -CV for this, but stress that it does not denote ‘ L -coefficient of variation’, saying that it is more properly described as a ‘coefficient of L -variation’. In this spirit we refer to it as the cLv where occasion arises.

4.2.3 L -moments and traditional moments

First, we note that L -moments have the following properties:

1) *Existence* If the mean of the distribution exists, all the L -moments exist. In contrast, existence of the higher moments is not guaranteed in the traditional system. For example, in the generalised extreme value distribution the third and fourth moments do not exist when the shape parameter k satisfies $k \geq 1/3$ and $k \geq 1/4$ respectively (the L -moment ratios are $\lambda_3 = 0.403$ and $\lambda_4 = 0.241$).

2) *Uniqueness* If the mean of the distribution exists, then the L -moments uniquely define the distribution; no two distributions have the same L -moments.

3) *Numerical values*: λ_1 can take any value; $\lambda_2 \geq 0$; L -moment ratios satisfy $|\lambda_r| < 1$ for all $r \geq 3$ (and tighter bounds can be found for individual λ_r).

Traditional population moments, and hence the associated skewness and kurtosis measures, are unbounded. Arguably the boundedness of the higher L -moment ratios is more intuitive, since it is coherent to think of a distribution having maximal kurtosis (indefinitely large density at a single point in the distribution's range and negligible density elsewhere) and maximal skewness (indefinitely large density at one end-point of the range, with zero density to one side and negligible density to the other).

Next, we compare the first four moments numerically. The first L -moment λ_1 , being the mean, is the same as the first ordinary moment μ_1 . In the case of λ_2 and the standard deviation σ (the product-moment-based scale measure) the two measures satisfy the inequality $\lambda_2 \leq \sqrt{3} \sigma^2$; many moderately-skewed distributions have $\lambda_2 \approx 2 \sigma^2$ (exactly so in the case of the exponential distribution). The CV and cLv bear a similar relationship; their estimators satisfy

$$\hat{C}_v = \frac{3n}{n+1} t^{\frac{1}{2}}$$

but for moderately skewed distributions \hat{C}_v may be twice as large as t , while if outliers are present \hat{C}_v is larger still.

To see why the traditional measure of dispersion must exceed the L -moment measure, note that in a sample of size 2

$$s^2 = \frac{1}{2} E(X_{2:2} - X_{1:2})^2, \quad L^2 = \frac{1}{2} E(X_{2:2} - X_{1:2})^2$$

Evidently L^2 gives more weight to larger differences – a point which applies to comparisons of the higher moments also, and is regarded by Hosking and Wallis (1997:38) as the main difference between the two systems. This reduced sensitivity to extreme values has a clear attraction in the face of the profit rate data summarised in the previous section.

In the cases of skewness and kurtosis comparison is subject to the point that the L -analogues are bounded while the standard versions are not; but one may note that the skewness of some distributions approaches infinity while the L -skewness is still very modest. (For example, graphing L_3 against β_3 for the generalised logistic distribution produces a curve whose slope already approaches the vertical at $\beta_3 = 0.3$, at which point $L_3 = 6$; see Hosking and Wallis 1997:36, Figure 2.11.)

But perhaps the most important advantage of the L -moment system relates to their respective sample statistics and the relations of those to the population statistics. Sample L -moment ratios can take any value which the population L -moment ratios can; by contrast in the traditional system the sample skewness g and kurtosis k are subject to algebraic bounds related to sample size

$$|g| \leq n^{1/2} \text{ and } k \leq n + 3$$

By way of example, a two-parameter log-normal distribution with $\beta_3 = 1$ has skewness 6.91, but samples of size 20 drawn from such a distribution cannot have sample skewness greater than 4.47, or 65 per cent of the population value (Hosking and Wallis 1997:18).

The result is that estimators of the traditional moments may be severely biased. On the other hand, although the L -moment ratio estimators t_r are not unbiased estimators of the corresponding statistics, the bias is very small in moderate or large samples (Hosking and Wallis 1997:28).

4.2.4 L -moment ratio diagrams

A convenient way of representing the L -moment ratios of different distributions is an L -moment ratio diagram such as Figure 4.1, which illustrates the L -skewness and L -kurtosis universe and plots the loci of a number of distributions in that space. The shaded area denotes combinations of L -skewness and L -kurtosis values which no distribution can have, and demonstrates that L -kurtosis not only has an upper bound of 1, but a lower bound of -0.25 . The polynomial approximations of t_4 in terms of t_3 used in the construction of the L -moment ratio diagram are given in Hosking and Wallis (1997:207).

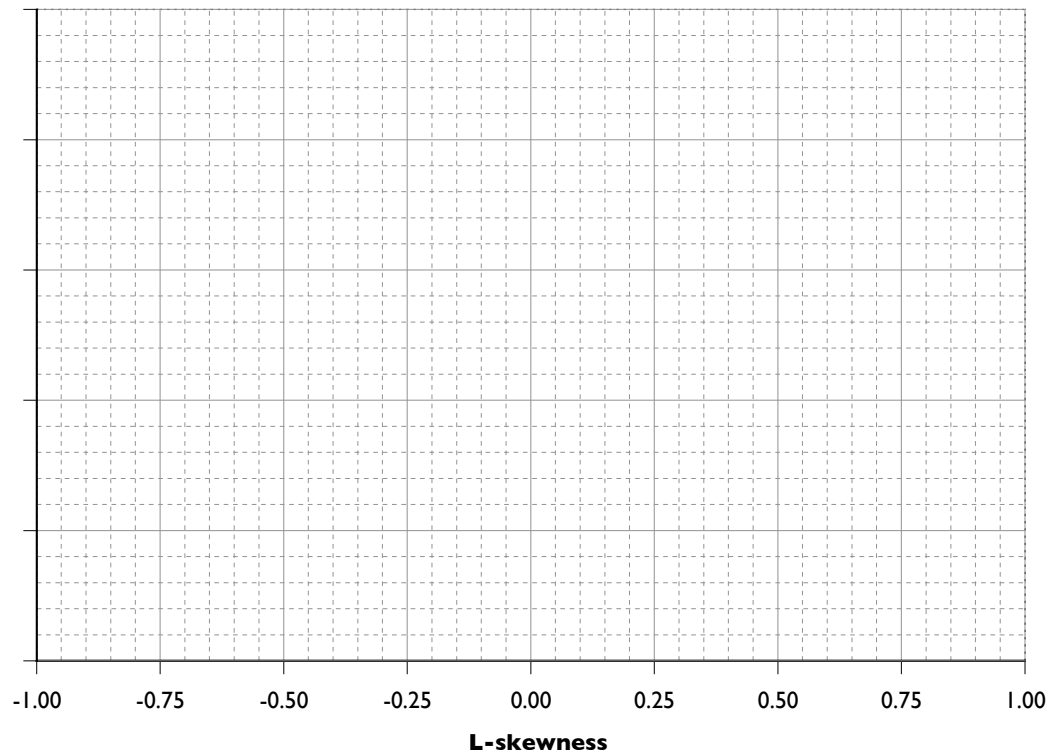


Figure 4.3: loci of selected distributions in L -skewness, L -kurtosis space (t_3 , t_4 space)

Two-parameter location-scale distributions appear as points because if the only differences between distributions X and Y are location and scale parameters then they are distributions of random variables X and $Y = aX + b$, and these random variables have the same L -skewness and L -kurtosis. Particular cases of note include the uniform (identified by ■ in the Figure), which with L -skewness and L -kurtosis both equal to zero plays a similar role in L -moment theory as the Gaussian (♦) does in cumulant theory; the exponential (●), which has the easily-memorised L -moment ratios $\lambda_3 = \frac{1}{3}$ and $\lambda_4 = \frac{1}{6}$, and the Gumbel (▲) (a special case of the generalised extreme value distribution).

Three-parameter location-scale-shape distributions have linear loci, with the position on the line corresponding to different shape parameters. Figure 4.1 shows six of these: considered at the $\lambda_3 = 0.25$ ordinate they are, from top to bottom: the generalised logistic (GLO); the GEV; the three-parameter lognormal (LN3); the three-parameter gamma, or Pearson Type III (PE3); the Weibull; and the generalised Pareto (GPA). The location of the exponential at the intersection of the gamma and Weibull reminds us not only that the exponential is a special case of both, but that all three are special cases of a generalised gamma distribution with two shape parameters (see Appendix, and McDonald, 1984).

As might be expected from the previous section, identification of the distribution from which an observed random sample was drawn is much more easily done using L -moments than with conventional moments. Hosking (1990) used Monte Carlo methods to show that conventional skewness and kurtosis were unable to discriminate among a generalised extreme value distribution and two different Weibull distributions. In contrast, the clouds of points in L -skewness, L -kurtosis space representing the repeated samples from the differing distributions were clearly distinguishable.

4.2.5 Parameter estimation using L -moment ratios

Hosking and Wallis (1997:191ff) present parameter estimators in terms of polynomial functions of L -moment ratios for the distributions whose loci in L -skewness, L -kurtosis space are illustrated in Figure 4.1. Estimation is extremely straightforward: for any given distribution (perhaps chosen on the basis of the data's location in Figure 4.1) first the

estimated shape parameter k is found as a function of s_3 , then the scale parameter in terms of k and s_2 , and finally the location parameter in terms of s_2 and s_1 .

4.2.6 Model selection using L -moment ratios

The L -moment system's origins are in the hydrological literature, as are its principal applications to date. Hydrologists need to provide answers to questions such as the following: 'How many years is it likely to be before the next flood of size x ?', where x denotes a number such that the flood described is not merely large, but disastrously so – in other words they are asked to produce accurate estimates of rare and extreme events, which requires the ability to apply distributional models which are accurate far into their tails and not merely in the main mass.

The desirable properties of L -moments described earlier mean that they can be reliably used to identify and estimate distributions, provided that sets of data generated by the same mechanism are available from a number of locations. In hydrology these will be gauges producing time-series data on the flow at a range of spatially-distributed points. In our case, we will be using cross-sectional data at temporally-distributed measuring points: in other words, our set of profit rates calculated from annual company accounts data.

4.2.7 L -moment summary

To summarise, the features of the L -moment system which make it particularly suited to the present study are:

- 1) dispersion and shape statistics are robust in the presence of extreme values
- 2) sample statistics may take any value which population statistics can achieve
- 3) parameter estimation by simple polynomial functions of sample L -statistics
- 4) from (1) and (2), more reliable model selection using sample statistics, especially when accuracy in the tails is important
- 5) from (1), (2) and (3), they are suited to method of moments estimation of distributional models.

4.2.8 Application of L -moments to profit rate data at the firm level

With the above discussion in mind, we now plot estimated L -moment ratios for the 21 profit rate measures which we defined in Chapter Two, Table 2.1, and constructed as summarised in Chapter Three, Tables 3.1 and 3.2). For each measure we use the data of Chapter Three to directly calculate the cLv , γ_3 , and γ_4 – the coefficient of L -variation, the L -skewness and L -kurtosis.

As can be seen from Figure 4.3, a number of our measures are *apparently* in a region where discrimination between models verges on the pointless. The word ‘apparently’ is stressed: if we refer back to the non-parametric density estimates of Figures 3.3 to 3.6, their main probability masses do not appear to be candidates for extreme skewness and kurtosis. Since resistance to the influence of extreme values is one of the most salient features of L -moments, it is clear that our profit rate measure distributions must include some very extreme values indeed.

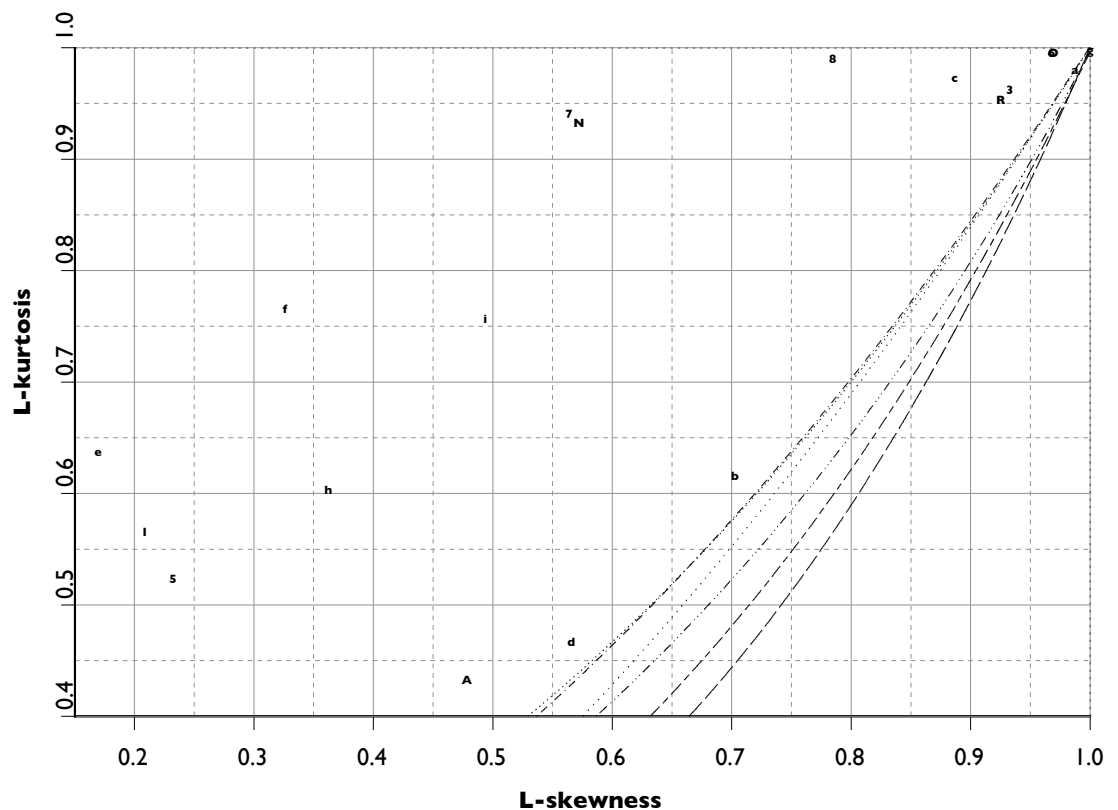


Figure 4.4: γ_3 , γ_4 space chart for 21 profit rate measures; see Table 4.1 for key

Table 4.1: key to Figure 4.4

Profit rate	Key letter	Profit rate	Key letter	Profit rate	Key letter
Gillman 1	a	ROCE	R	Glick 1	1
Gillman 2	b	ORE	O	Glick 2	2
Gillman 3	c	ATO	A	Glick 3	3
Gillman 4	d	NPM	N	Glick 4	4
Gillman 5f	e			Glick 5	5
Gillman 5s	f			Glick 6	6
Gillman 6	g			Glick 7	7
Gillman 7f	h			Glick 8	8
Gillman 7f	i				

4.3 Estimating capital-level profit rate distributions

4.3.1 Randomly-sized random samples

So far we have considered the distribution of profit rate measures across the population of firms, and this would be appropriate in testing Gibrat's log-normal hypothesis. But a test of Farjoun and Machover's gamma distribution hypothesis requires us to estimate the rate of return achieved by each unit of *capital* (which we take to mean each £1 of capital invested).

In principle we could look at the population distribution directly; we could take each observed value for a given profit rate measure, look at the size k_i of the firm concerned, as measured by the appropriate capital definition, and replicate the particular rate of return a further $k_i - 1$ times. But this would be computationally intractable, given the enormous number of observations that would result – one for each of more than 1,284bn £1 units of capital, in the case of Gillman 1, for example.

Thus some method of sampling will be preferable, which immediately raises the question of how big a sample to take. This is a difficulty. For example, the standard goodness-of-fit test is the Kolmogorov-Smirnov (KS).⁴¹ This essentially compares the empirical cumulative distribution function to that of the of the estimated model and assesses whether the

⁴¹ The Kolmogorov-Smirnov test is possibly the best-known, but alternatives are the Cramér-von Mises and Anderson-Darling tests (for further references, see Anderson and Darling, 1952).

maximum vertical distance between the two can be thought of as arising by chance. Informally, the more data one has, the greater the likelihood of noise giving rise to rejection of what might otherwise be a well-fitting model. However, one clearly cannot choose the size of sample with an eye on whether it makes it more (or less) likely that any particular model will pass the KS test. Happily, it turns out that this invidious choice can be avoided, thanks to a scheme suggested by Chris Jones (personal communication).

Suppose that the profit-rates p_i of each of i firms, according to a given profit rate measure, are entered once each in a draw, in which the probability of being chosen is k_i/k_{\max} , where k_i is the size of the i -th firm according to the capital measure used in the profit rate measure, and k_{\max} is the size of the largest firm. Clearly the largest firm is guaranteed to have its profit rate p_{\max} selected, while the probability of other observed profit-rates being drawn will be in accordance with the relevant firms' size compared to that of the largest. However, the total *number* of profit-rates drawn will be random – we will have a randomly-sized random sample (RS2).

Such an RS2 is an unbiased estimator of the overall population and hence the sample value of any statistic of interest should be an unbiased estimator of the population statistic. To see this, consider conducting m such draws and concatenating the results: p_{\max} will naturally appear exactly m times, while the remaining p_i will tend to appear mk_i/k_{\max} times. As $m \rightarrow k_{\max}$ each profit rate will, in probability, appear k_i times, and p_{\max} exactly k_{\max} times, giving a close approximation to the distribution of profit-rates across the underlying population of capital units (for example, £1s), randomly divided among the m samples. The individual sample statistics will of course be subject to variance but, given standard assumptions about their distribution converging to the Gaussian, averaging the sample values should give a result close to the population value. We will revisit this assumption in Chapter Five.

We could, of course, actually take k_{\max} RS2s, estimate any desired statistic as just suggested, and cross-check against the value for the simulated population. However, given the actual value of k_{\max} according to the various capital measures, this would require

unfeasible amounts of computation. The question is thus how few RS2s can we take while securing adequate protection against variance. The procedure outlined is akin to the bootstrap procedure (Efron and Tibshirani, 1993), and a rule of thumb for the bootstrap is that for datasets of ‘reasonable’ size, 100 replications should be enough.

It turns out that because of the extreme range and skewness of k_i under all the profit rate measures (the vast majority of firms are very small), the RS2s drawn as described are rather sparse (in two cases – Glick 2 and 4 – sufficiently sparse that it is not guaranteed that a given RS2 will be large enough to calculate $\hat{\alpha}_i$; in other words, some of these samples are smaller than four). The solution, however, is simple: treat the concatenation of the m samples as a single larger RS2 sample and calculate the desired statistic, replicate the modified RS2 procedure n times, and take an average of the statistics. In our procedure $n = m = 100$, and thus each profit rate estimate involves 10,000 separate random samples.

We note that the above also implements the fuzzy rejection strategy we suggested for coping with contaminants in a firm-level test.

4.3.2 Empirical densities and summary statistics for capital-level data

In this section we present an exploratory analysis of the capital-weighted profit rate measure distributions on the same pattern as for the firm-level distributions in Chapter Three. We thus begin with empirical density functions estimated by histograms. Figure 4.5 shows a detailed histogram for the metasample (that is, the concatenation of 100 RS2 samples) of Gillman 1, together with a kernel density estimate (KDE).

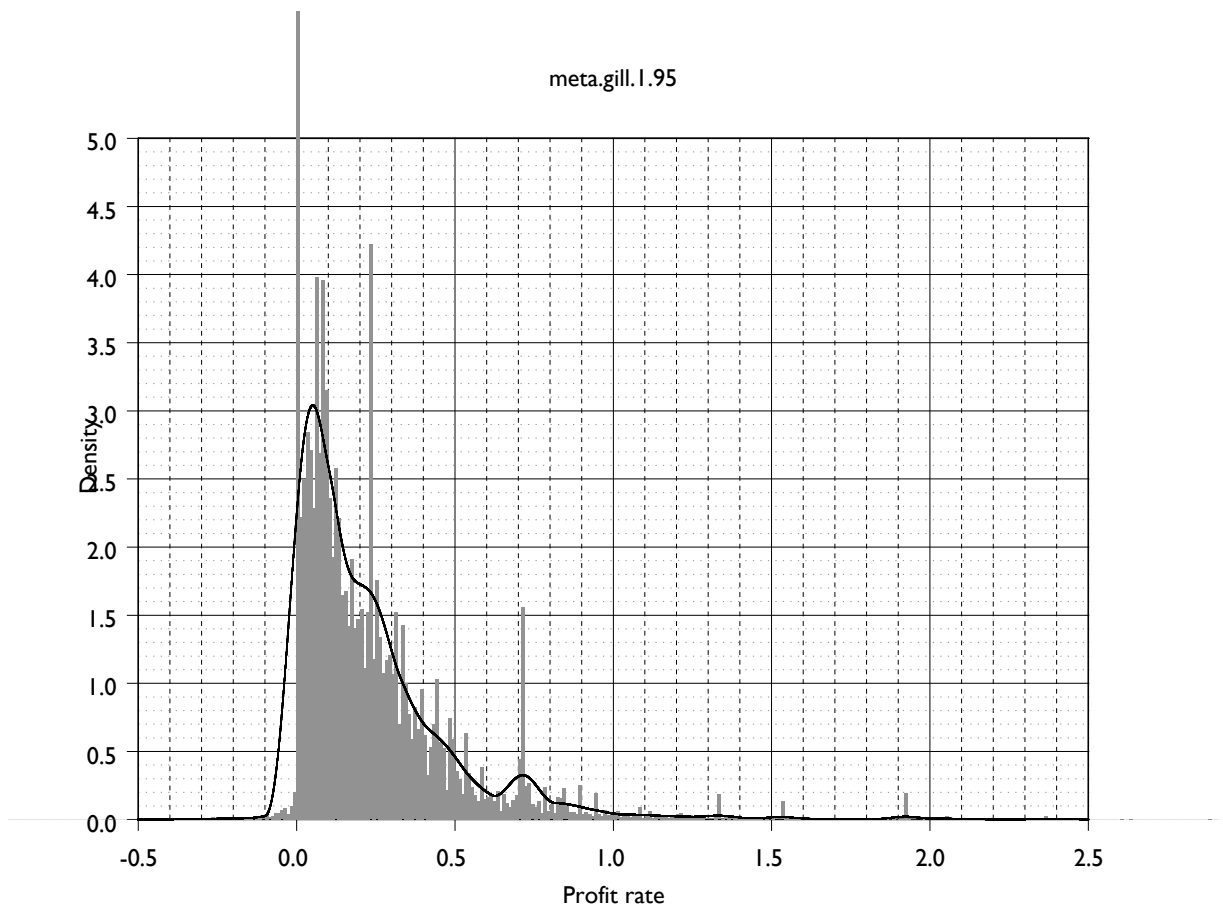


Figure 4.5: empirical density estimate for Gillman I, metasample of weighted data

An obvious difference is the increased noisiness of the histogram estimate. This is evidently because of the high degree of concentration of firm size – note that the full height of the bar on the 0, 0.01 interval is in fact 10.5, in other words this one interval accounts for approximately 10 per cent of the total capital, while the largest firm accounts for 5.6 per cent of the whole sample (10,000 occurrences in a sample of 179,614). This extra noise raises the question of what bandwidth to use for the KDE estimate. Here we have used $h = 0.2$, modified by the appropriate factor for choice of kernel (in other words twice the bandwidth used in the equivalent Figure 3.2 above). This choice is made using the same *ad hoc* criterion as before, namely that it provides ‘adequate’ smoothing (and is a simple multiple of the value of h used above).

The question of smoothing has rather more significance here than above, and not only because of the higher level of noise. The histogram depends on the assumption that all units of capital employed by each firm in fact earn the same rate of return – an assumption

which becomes increasingly unlikely with firm size. Applying any of the common kernels amounts to asserting a definite claim about within-firm distribution of the profit rate measure, in particular that it is identical in scale for all firms (and moreover symmetrical). Both these features seem unlikely. For the largest firms, probably highly integrated both vertically and horizontally, it would be more plausible to argue that the within-firm distribution closely resembles that of the economy as a whole, perhaps with a change of location parameter. But since the economy-wide distribution is precisely what we are attempting to discover, this is not a hypothesis that can easily be made operational, though one might envisage some kind of iterative procedure. Moreover, since the theoretically-predicted distributions are asymmetrical, this implies the use of non-symmetrical kernels. There seems to be little literature on these, although one simple approach would be to transform the data to approximate normality, smooth with a symmetric kernel, and transform back.

A further important point concerns the tails of the weighted distribution. Close inspection of Figure 4.5 reveals small spikes at around 1.5, 1.7, 1.9 and 2.2. Are these outliers, or signs that the heavy tails seen in the data in Chapter Three have not been wholly eliminated by the RS2 weighting process?

Comparative histograms for each weighted profit rate measure are given in Figures 4.6 to 4.9. Once again, limitations of reproduction at this scale mean that some fine detail is lost, especially where there are large spikes as in evident in Figure 4.5. However, in some cases there is clear evidence that heavy tails may still be present. Consider, for example, Gillman's 'marxian' measures in Figure 4.6, and in particular Gillman 3 and 4, or the histogram of ATO in Figure 4.8. Also compare Figure 4.5 with its counterpart in Figure 4.6, the top left panel: the larger version shows small spikes at around 1.5, 1.7, 1.9 and 2.2 which fail to be reproduced in the smaller version. In fact these are a feature of the histograms of all measures, even though they do not always show clearly in the versions reproduced here.

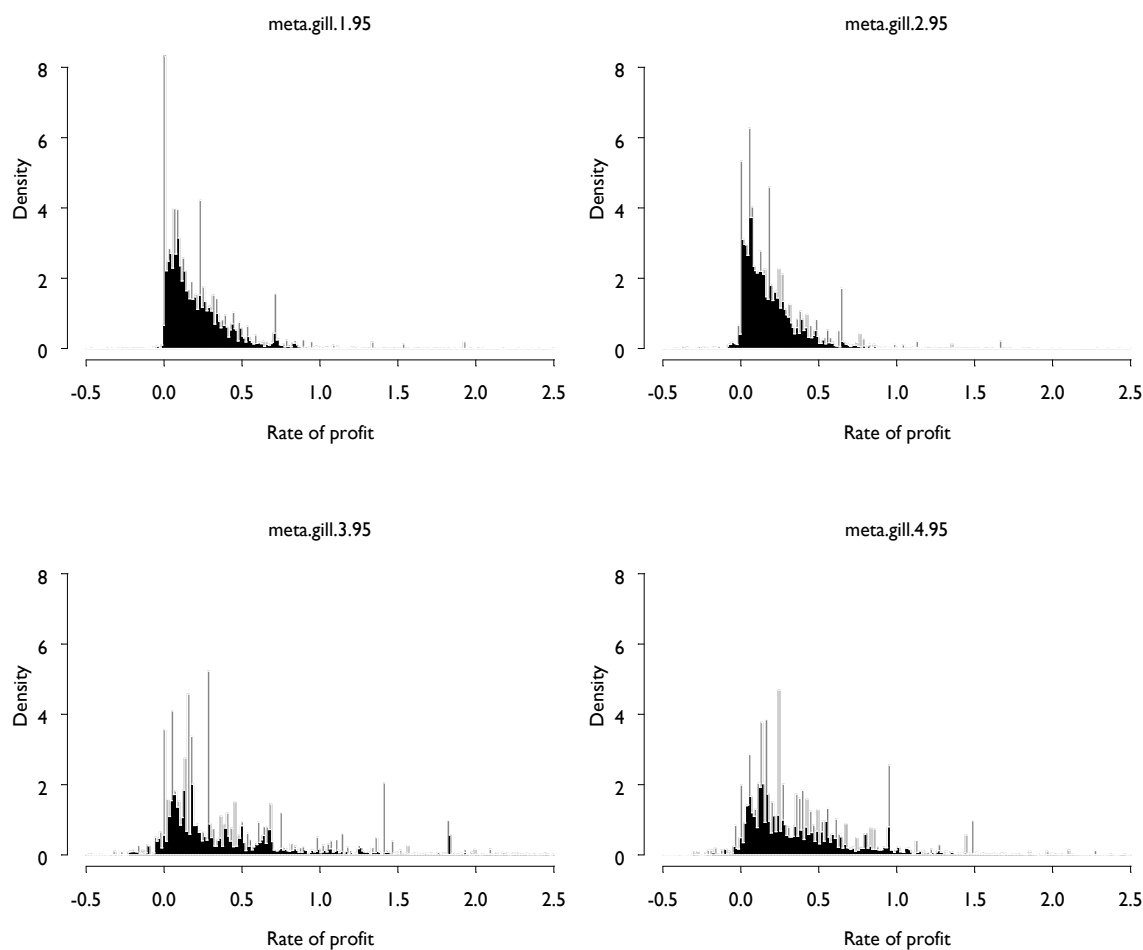


Figure 4.6: empirical density estimate for Gillman's 'marxian' measures, metasample of weighted data

Nonetheless, it appears that the capital-weighting procedure has not only reduced the range of all the measures (seen in Figure 4.10 below), but also induced changes in the shapes of many distributions, especially those which had less peaked shapes at the firm level. In particular, Gillman 1 to Gillman 4, the four financial accounting measures, and all the Glick measures, with the exception of Glick 7, exhibit marked changes (compare Figures 3.3 to 3.6).

In the case of the Gillman measures, the effect is to make them look very strong candidates for an exponential model (or near-exponential, if the KDE in Figure 4.5 is borne in mind).

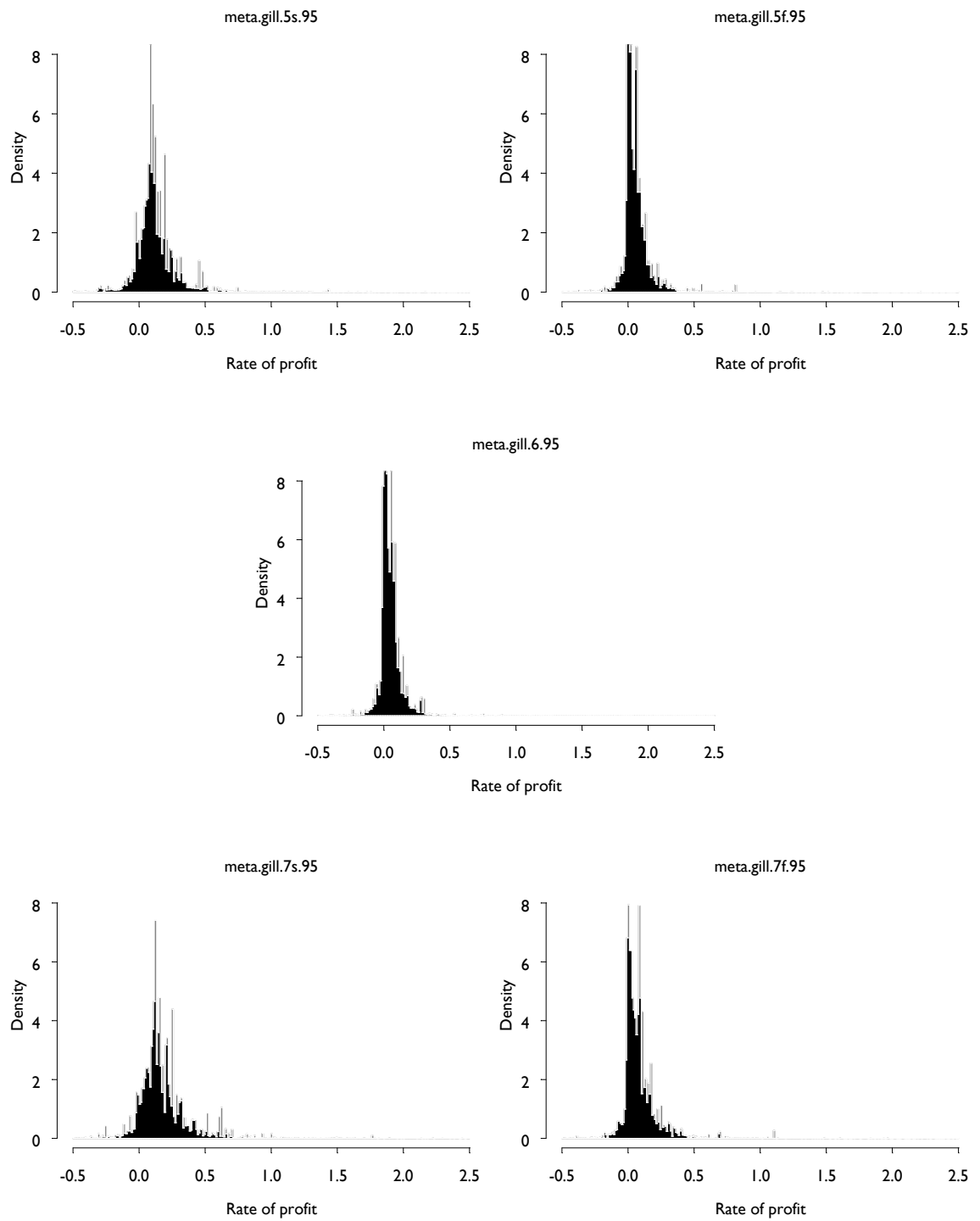


Figure 4.7: empirical density estimate for Gillman's 'capitalist' measures, metasample of weighted data

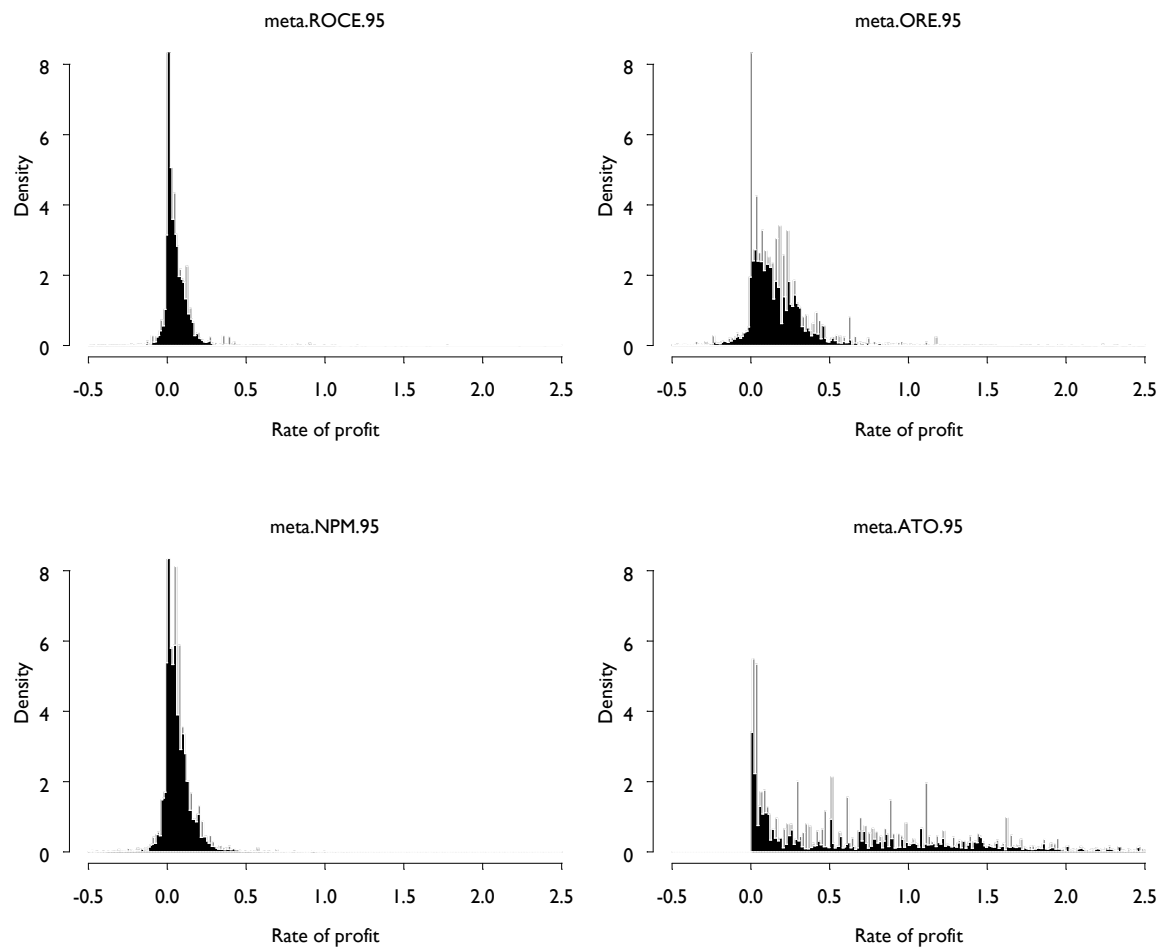


Figure 4.8: empirical density estimate for accounting ratios, metasample of weighted data

A further point to note is that the general similarity between the three accounting ratios ROCE, ORE and NPM and the Glick profit rate measures Glick 1, Glick 6, and Glick 7 is broadly preserved: see Figures 4.8 and 4.9, and compare with Figures 3.5 and 3.6.

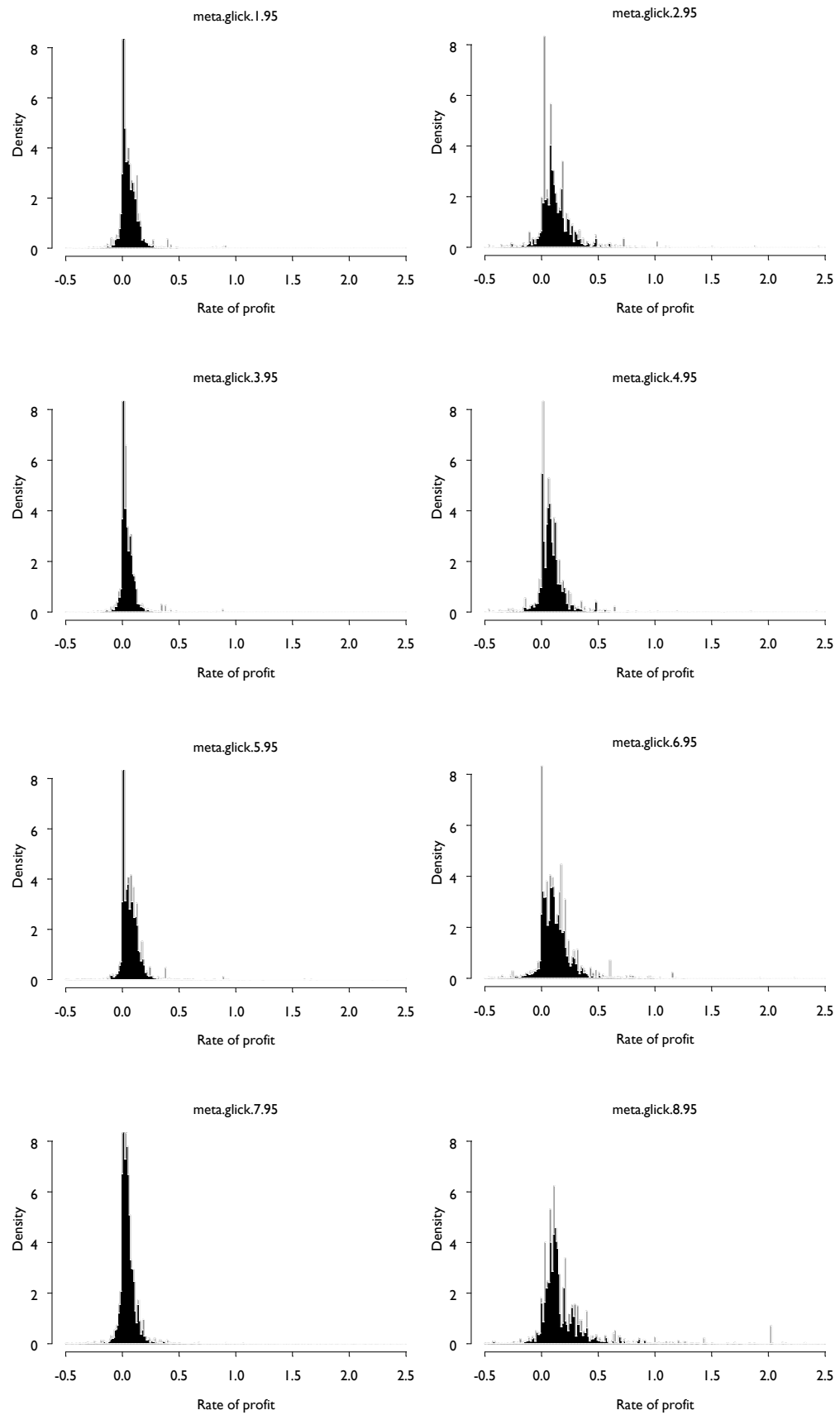


Figure 4.9: empirical density estimate for Glick measures, metasample of weighted data

Table 4.2 gives estimates of several measures of central tendency, and of the skewness, kurtosis and range of our weighted profit rate measures; the method is, as discussed above, to compute the relevant statistics for each RS2 sample and to average the results (we also show the average RS2 sample size).

That the RS2 procedure is capable of giving appropriate results can be seen immediately, by comparing the general rate of profit (GRP) in each case (repeated from Table 3.3 above) with the estimated mean for each measure. Since the GRP is by definition the average rate of profit per unit of capital, any estimate of the capital-weighted mean should be close to the GRP – as indeed it is in nearly all cases: generally within one percentage point, with the exceptions being mainly where the GRP is high to begin with.

A further interesting point concerns the estimated modes. Here the individual sample modes were estimated using cosine kernels with bandwidth 0.1, just as was done in the firm-level case. Yet the averaging process has produced an estimate – for Gillman 1 – which is evidently close to the one-off estimate for the meta-sample mode, as found in Figure 4.5 by inspection of the KDE.

As might be expected from the histograms, the conventional skewness and kurtosis are markedly reduced in all cases, though they are still large (compare with Table 3.3). The most striking result here is that several profit rate measures now show negative skewness, mostly of a relatively mild degree except for Glick 6 (-26.48).

Table 4.2 (a): summary statistics for firm-weighted profit rate distributions, Gillman's 'marxian' measures

	gill.1	gill.2	gill.3	gill.4
ave.obs	1796.14	3946.68	2253.72	2497.48
GRP	0.2451	0.2162	0.5972	0.4518
Mean	0.2481	0.2161	0.6252	0.4735
median	0.1649	0.1478	0.2838	0.2823
Mode	0.0647	0.0754	0.1285	0.1242
Mean.10	0.1921	0.1741	0.4001	0.3660
Skew	5.2978	7.5072	21.0492	24.7939
Kurt	51.2628	134.80	629.76	932.32
Cv	1.3339	1.6394	2.9885	1.9874
Mad	0.1838	0.1522	0.3348	0.2823

Table 4.2 (b): summary statistics for firm-weighted profit rate distributions, Gillman's 'capitalist measures

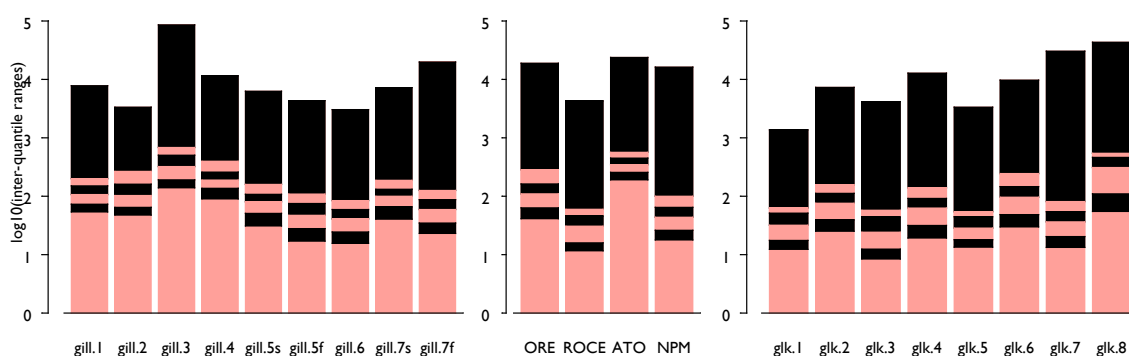
	gill.5s	gill.5f	gill.6	gill.7s	gill.7f
Ave.obs	2500.09	3931.68	4107.07	2285.69	3678.59
GRP	0.1232	0.0594	0.0526	0.1844	0.0880
Mean	0.1261	0.0598	0.0528	0.1831	0.0863
median	0.1049	0.0453	0.0398	0.1490	0.0603
Mode	0.0933	0.0283	0.0287	0.1104	0.0378
Mean.10	0.1160	0.0487	0.0452	0.1611	0.0676
Skew	0.4157	-0.7115	1.1640	4.5661	-2.2736
Kurt	34.2991	327.22	28.54	66.11	166.88
Cv	1.6078	2.5366	1.9645	1.4706	2.1177
Mad	0.0960	0.0589	0.0538	0.1158	0.0746

Table 4.2 (c): summary statistics for firm-weighted profit rate distributions, accounting ratios

	ROCE	ORE	ATO	NPM
Ave.obs	3267.82	2525.37	2377.23	3967.32
GRP	0.1756	0.0387	0.8903	0.0700
Mean	0.0686	2.6307	0.8985	0.0676
median	0.0475	0.2592	0.6138	0.0489
Mode	0.0342	0.1871	0.0606	0.0315
Mean.10	0.0557	0.2499	0.7249	0.0544
Skew	1.3139	37.02	7.0525	19.58
Kurt	48.53	1545.10	113.04	753.59
Cv	1.9113	22.36	1.3024	3.2262
Mad	0.0586	0.1915	0.8004	0.0617

Table 4.2 (d): summary statistics for firm-weighted profit rate distributions, Glick measures

	Glick.1	glick.2	glick.3	glick.4	glick.5	Glick.6	glick.7	glick.8
ave.obs	1690.95	683.88	2527.28	714.19	1430.07	3402.63	3966.42	3540.16
GRP	0.0429	0.1096	0.0294	0.0794	0.0517	0.1339	0.0500	0.3364
Mean	0.0409	0.1162	0.0301	0.0757	0.0511	0.12614	0.0464	0.3169
median	0.0162	0.0720	0.0092	0.0471	0.0166	0.096641	0.0312	0.1307
Mode	0.0263	0.0625	0.0145	0.0506	0.0338	0.05475	0.0223	0.1062
Mean.10	0.0328	0.0898	0.0205	0.0588	0.0421	0.10551	0.0367	0.1670
Skew	0.4348	-0.979	-1.2685	3.5260	-0.6105	-26.4783	24.21	14.28
Kurt	83.26	50.88	101.12	32.12	23.46	1284.273	890.58	319.92
Cv	2.7302	1.9629	2.9704	2.0417	1.6960	3.478679	3.2067	3.2338
Mad	0.0214	0.0732	0.0107	0.0503	0.0263	0.11824	0.0425	0.1121

**Figure 4.10: log ranges of capital-weighted profit rate distributions**

Finally, in Figure 4.10 we produce the counterpart to Figure 3.7, the analysis of the extreme tails of the distributions (again, for the meta-samples for each profit rate measure). These are concatenations of 10,000 samples, and thus might be expected to include some quite extreme values by chance, even given our RS2 weighting scheme. Nonetheless, not only are the pathologically large ranges associated with Gillman 6 and Glick 2 and Glick 4 no longer present, but the ranges of all the measures are much reduced, to the extent that the largest tick mark on the (log-scaled) vertical axis is now 5 (that is, a range of 10^5 per cent), whereas before it was 8 (a range of 10^8 per cent).

Further, the relative lengths of the extreme tails has been reduced, even though the range of the central parts of each distribution has been reduced (the range of the central 80 per cent of Gillman 1 has fallen from just above to noticeably below 100 per cent).

4.4 Model selection with capital-level data

In this section we estimate L -moment ratios for each of our 21 profit rate measures and plot their positions in β_3, β_4 space, with a view to gaining insight into possible distributional models for later estimation. Actual selection of models requires clouds of points from annual data; this will be illustrated in Chapter Five.

To estimate the L -moment ratios for each of our 21 profit rate measures we therefore use the RS2 procedure described above with $m = n = 100$ to estimate $\hat{\beta}_i$, where $i = 1, 2, 3, 4$, and use the resulting estimated L -moments to calculate the L -moment ratios $\hat{\beta}_i$.

The results of this are shown in Figure 4.11. The latter should be compared with Figure 4.2, which illustrated the L -moment ratios for the firm-level data; the difference between the two is the result of the application of our RS2 sampling scheme to estimate L -moment ratios at the capital level.

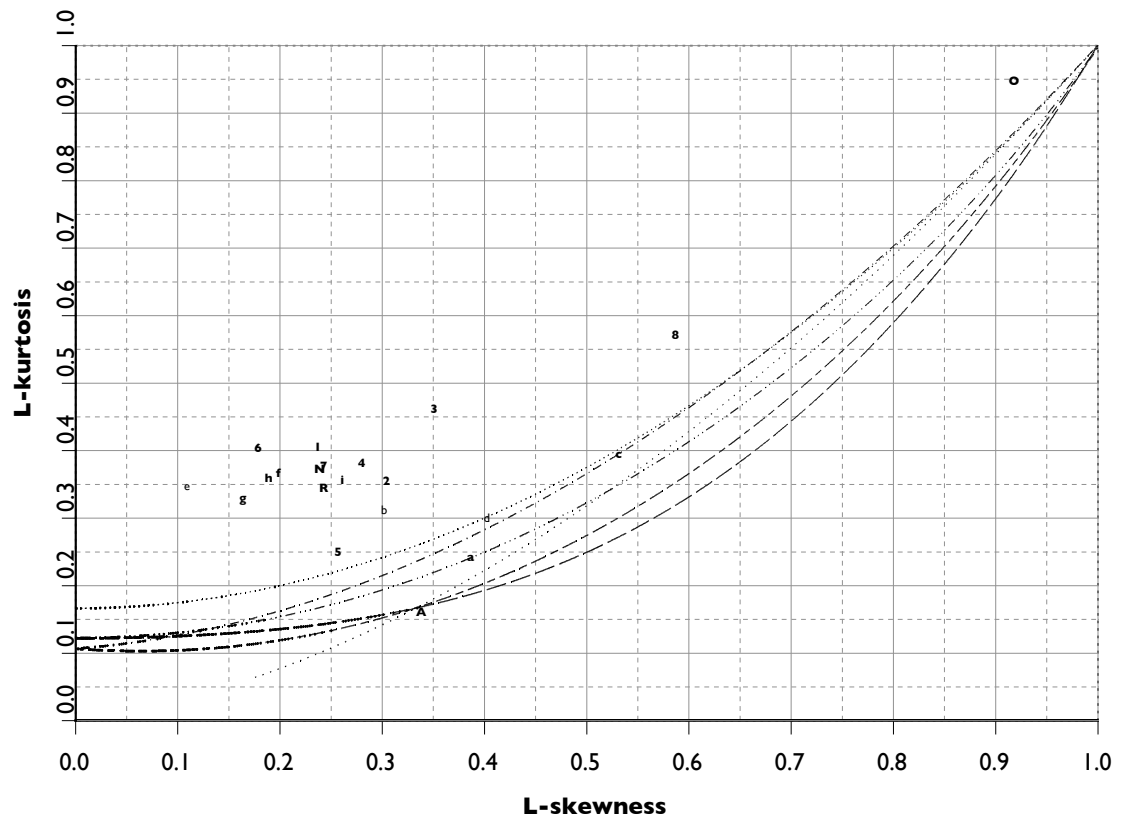


Figure 4.11: β_3, β_4 space chart for weighted profit rate measures; see Table 4.1 for key

Nearly all the measures now have sharply reduced skewness and kurtosis (although ORE and Glick 8 are exceptions – points ‘O’ and ‘8’ on the plot). On the other hand, most of the measures lie well above the band of distributional loci provided by Hosking and Wallis (1997).

More encouragingly, three of Gillman’s ‘marxian’ measures *do* lie in this band – Gillman 1, 3 and 4 (points ‘a’, ‘c’ and ‘d’ on the plot), while Gillman 2 (point ‘b’) is relatively close.

4.5 Conclusion

In this chapter we have addressed a number of problematic features revealed by the exploratory data analysis of Chapter Three, where we found that all profit rate measures had very extended tails, in most cases to the left as well as to the right of the main body of data.

To deal with these discordant tails we proposed the use of the method of *L*-moments (to our knowledge, for the first time in the economics literature). This was not only because of their robustness in the face of outliers, but because they offer a straightforward way of estimating threshold parameters for a number of distributions with bounded support. However, their direct application to the firm-level data showed that even this robust system of moments found extraordinary levels of *L*-skewness and kurtosis.

We also investigated the use of randomly-sized random samples to estimate *L*-moment ratios. This too had a double objective. First, we conjectured that the many extreme values found in the company-level data represented very small, non-capitalist enterprises: weighting observations by size of firm amounts to a scheme of fuzzy rejection of these presumed contaminants. Secondly, we in any case need to weight the data by firm size in order to test Farjoun and Machover’s conjecture about the profit rate distribution, which applies to the capital level, not the firm level.

The RS2 procedure reduced the extent of the tails, as shown by charting the inter-quartile ranges for the different profit rate measures, while histograms show very differently-shaped distributions to those for the original firm-level data.

The L -moment ratios estimated from the RS2 samples showed marked reductions in skewness and kurtosis. Plotting the results for the 1995 data for the various profit rate measures in L -skewness and kurtosis space produced mixed results, but these were encouraging from the point of view of testing both Gibrat's and Farjoun and Machover's hypotheses, since the four Gillman 'marxian' measures appear to be good candidates for modelling by exponential distributions for which Hosking and Wallis (1997) have supplied methods.

In Chapter Five we will make full use of our combined RS2 and L -moments methods (RS2- L) by plotting several years' data as clouds of points in L -skewness and kurtosis space so as to more precisely identify candidate distributions.

Chapter 5 The distribution of the rate of profit

In this chapter we test the ability of the techniques developed in Chapter Four to identify the distributional laws followed by different measures of the rate of return. It is thus a test of the hypotheses of Gibrat (1931) and of Farjoun and Machover (1983).

Since our method involves inspecting clouds of points in L -skewness, L -kurtosis space which represent the profit rate in different years, it is simultaneously an exploration of how the precise forms of the distributions vary over the business cycle (but a necessarily limited one, in view of the small number of years covered by our data).

Gibrat's hypothesis of a log normal distribution of the profit rate is simply one of many illustrative examples intended to demonstrate the wide applicability of his Law of Proportionate Effect (LPE).⁴² As such, it is not backed by arguments specific to the rate of profit; indeed, we saw in Chapter Two that Gibrat declines to define profit himself (1931:180), although his discussion draws on previous studies of dividends. Moreover, it appears that no one has followed up his suggestion about profit rates.

Likewise, there has been limited work on Farjoun and Machover's notion that the profit rate is a random variable, even though their small following is extremely respectful of their work.⁴³

⁴² Other economic quantities for which Gibrat provides illustrations include not only the celebrated study of firm size, but also income, wealth, and size of towns. Note that despite the indelible association of this law with Gibrat in the economics literature, similar processes were described by Kapteyn (1903) in relation to biological statistics, while the log normal itself was first described by McAlister (1879).

⁴³ A number of workers cite Farjoun and Machover's approach with respect, but do not test their hypotheses directly (Cockshott and Cottrell, 1994, 1998, Cockshott et al., 1995, Julius, 2005, Puty, n.d., Sheppard and Barnes, 1986, 1990, Webber, Michael, 1987, Webber, Michael J. and Rigby, 1986, Webber, Michael J. and Rigby, 1996). A partial exception is Zachariah (2005); but although he estimates empirical probability density functions he does so for *industries*, not firms, and does not attempt to identify any functional form for them, still less estimate parameters. Wright's work (2005) cited above is part of a project to build an agent-based simulation model of a capitalist economy in which the outcomes of transactions are random variables and the limiting distributions of their outcomes (including the profit rate) conform to accepted stylised facts.

However, Farjoun and Machover's stipulations about the definition of the profit rate, and about the population to be considered, are much more specific than Gibrat's. Moreover they provide an elaborate theoretical justification for their hypothesis. Indeed, that hypothesis is offered as just one confirming instance of the validity of a wide-ranging methodological outlook. Since the key objective of the current chapter is to test the Farjoun and Machover hypothesis, we necessarily devote some space here to elaborating the outline given in previous chapters.

Since Gibrat's case for the log normal distribution of profit rates is at best lightly argued in his book, we can conveniently state the LPE here before continuing: suppose some attribute of the members of a population is described by a variable whose value, for each particular member of the population, changes from one period to another according to some rule which is independent of the initial value for each member, but subject to random variation between members. It can be shown that the limiting form of the variable's distribution, after n rounds of change as n grows large, is the log normal distribution. For example, if over the interval t the size of each of i firms varies by $ak_i + e_i$, where a is some constant, k represents the size of each firm and e is a random variable, then in the limit the firm size distribution becomes log normal.

Since his illustrations take firms as the unit of observation, one might conclude that a strict confirmation of Gibrat would entail finding a log normal distribution at the firm level. We think that this would be unduly restrictive, and will take the detection of a log normal distribution at either firm or capital level, for any profit rate measure, as constituting support for Gibrat (in this study we only investigate functional forms for the capital level distribution).

Farjoun and Machover (1983) was not widely reviewed, but the notices which did appear included some by very significant figures (Champernowne, 1985, Steedman, 1983). The most penetratingly critical account was also the most sympathetic (Webber, M.J., 1986); a further review was by Jalali-Naini (1986).

The structure of the chapter is as follows: section 1 reviews Farjoun and Machover's probabilistic political economy and consider the implications of possible outcomes of our tests of their profit rate hypothesis. In section 2 we present charts of the annual estimated L -moment ratios for each profit rate measure, and consider the properties of our technique of randomly-sized random samples (RS2) in respect of these; section 3 renews our investigation of the tails of both unweighted and size-weighted versions of our profit rate measures.

5.1 Farjoun and Machover's probabilistic political economy

In this section we review Farjoun and Machover's case for adapting the statistical outlook of ideal gas mechanics to Marxian economics, and consider the possible outcomes of a test of their hypothesis about the profit rate distribution.

In one respect we diverge from their account: where they see their approach as one of 'restoring' the insights of *Capital* Volume I, lost through an unfortunate detour in Volume III as a result of the needless assumption of uniform profit rates, we suggest a different interpretation of Marx, arguing that he adopts a fundamentally probabilistic standpoint from the very beginning of his work in political economy through to its mature expression in *Capital*.

Finally, we discuss some practical considerations in testing Farjoun and Machover's hypothesis of a gamma distribution for the rate of profit, and the implications of the various possible outcomes.

5.1.1 Statistical mechanics, the transformation problem, and the gamma distribution

Our account of the statistical mechanics outlook, and of its applicability to issues to do with the rate of profit, substantially follows Farjoun and Machover (1983). A modern industrial economy contains a very large number of agents: in the United Kingdom, for example, there are tens of thousands of capitalist enterprises employing tens of millions of workers, and the commodities produced are the subject of billions of transactions every week.

Because the U.K. is a competitive capitalist economy, it is intrinsic that all this activity is unplanned, in the sense of not being directed by a central authority.

Since at least the time of Adam Smith, it has been a commonplace that all these activities, although unplanned, are nonetheless co-ordinated, and that this is a fact requiring explanation. Traditionally this explanation has been provided by reasoning about so-called representative agents⁴⁴ in terms of how their interaction results in the formation of ‘the’ price of a commodity, ‘the’ rate of profit, and so on, where the definite article implies a single uniform value for the variable concerned.

Of course, everyone knows that in practice economic variables do not have uniform values. The price of identical tomatoes in a street market may vary from stall to stall, while there would be little reason for the existence of a stock market if all companies achieved the same rate of profit. But it is assumed that the forces of competition will tend to smooth out such differences, given enough time and an absence of external disturbances, until uniform values are achieved in equilibrium. Indeed, uniformity is taken to be part of the meaning of equilibrium.

This notion of equilibrium is what Farjoun and Machover contest. As they point out, in Marx’s economics the forces that tend to bring about uniformity are opposed by other forces that tend to disrupt it. In the case of the rate of profit, movement of capital from one sector to another tends to equalise it, while the search for relative surplus value through technical innovation tends to differentiate it.⁴⁵ A firm which can reduce necessary labour will increase its profit rate, at least until such time as its competitors are able to copy the new techniques (or devise even better ones).

⁴⁴ The agents may be producers and consumers, or workers and capitalists, according to the theoretical tradition in question. But the fundamental method is widespread in all traditions.

⁴⁵ Relative surplus value is Marx’s term for the extra surplus value gained when capital is able to shorten the time needed for labour to reproduce its wage (necessary labour) within a *fixed* working day; absolute surplus value can be increased by lengthening the total working day while necessary labour time stays constant.

Both these forces are results of capitalist competition, hence internal to the system since competition is essential to the notion of capitalism. An adequate concept of capitalist equilibrium might therefore be one which accords equal status to both equalising and differentiating forces.

Taking uniformity of profit rates as the sole definition of equilibrium in effect treats competition as a force internal to the economic system insofar as it tends to equalise the rate of profit, and as an external, perturbing, force when it has the opposite tendency. As Farjoun and Machover point out (1983:36), this is like regarding the force of gravity as a force internal to a pendulum's system as it swings to the right (and gravity tends to pull it back to the left), but as an external force as it swings to the left (and gravity pulls it to the right).

A concept of equilibrium might instead be consistent in treating competition as an inherent feature of capitalism by incorporating both equalisation and differentiation, hence in treating it as an internal force. It should allow a range of profit rates to exist at any given moment, and be dynamic, in the sense of allowing individual firms' profit rates to vary through time. The required concept would be an equilibrium one in the sense that the proportion of the total social capital that achieves any particular rate of profit is roughly constant through time.

Here is the point of application of (classical) statistical mechanics: in an ideal gas, the velocities of each of the individual particles which compose it are taken to be both widely differentiated and rapidly changing as a result of constant interaction between them, in the form of collisions. The velocities, hence speeds, of the particles determine the heat energy of the gas, which is the sum total of the kinetic energy. This macroscopic variable is stable (assuming no input from outside the system). The law of the conservation of energy stipulates that if one particle gains energy from a particular collision the other particle must lose it. In a gas at macroscopic equilibrium (that is, neither heating up nor cooling down) the particles are in a sense competing for a share of a fixed pool of energy.

There is a clear analogy here with Marx's discussion in Volume III of *Capital*, where capitalists compete for profits in the form of an equal 'aliquot' share⁴⁶ in the total surplus value created in production. The overall rate of profit (compare: total heat energy) is fixed; the individual rates of profit are the result of individual competitive effort – in the form of both shifting capital from one industry to another, and investments in new technology – but the relative success of one firm is at the expense of the relative failure of another.

Farjoun and Machover note that the 'most chaotic' partition of kinetic energy among particles results in a gamma distribution (page 68), but that a formal proof would require a notion of entropy in the firm space.^{47,48}

The point of the analogy, for Farjoun and Machover (1983), is that assuming uniform speeds for the particles will give the wrong value for the total kinetic energy, essentially because the energy of each particle is proportional to the square of its speed. Thus the total energy is given by $nk \mathbf{E}(V^2)$, where n is the number of particles, $k = 0.5$ mass of each particle, V is the velocity of each particle, and $\mathbf{E}(V^2)$ denotes the mean of the squared velocities; assuming equal speeds would suggest that $nk (\mathbf{E}V)^2$ would also compute the total energy. But in general these expressions do not have the same value: $\mathbf{E}(V^2) - (\mathbf{E}V)^2$ is the standard definition of the variance of V .

Statistical mechanics uses probabilistic assumptions to derive the equilibrium distributions (and hence the expected values) of position, speed and energy among the particles at a given moment (Farjoun and Machover 1983:55). These are *space* averages (in the sense of sample space). But one can also consider the distribution of a given particle's speed, *etc.*, in *time*. To say that a system is in (dynamic) equilibrium is to say that, provided

⁴⁶ That is, an equal share for each unit of capital they have advanced.

⁴⁷ In fact Maxwell's distribution is a particular member of the χ^2 family, itself a special case of the gamma distribution (see Appendix).

⁴⁸ The notion of entropy is first introduced in a note to page 26, where Farjoun and Machover draw attention to another, independent, speculation on the possibility of a probabilistic economics (Jaynes, 1991).

that one considers a sufficiently long time interval, the time average approaches the space average as a limit.⁴⁹ An implication is that the time averages of two distinct particles will be approximately equal (Farjoun and Machover 1983:49).

As discussed in Chapter One, Farjoun and Machover (1983) is an intervention in the debate on the transformation problem: how to find a way of reconciling values determined by labour time with the prices needed to equalise the rate of profit between firms employing different proportions of capital and labour. Their ‘*dissolution*’ of the transformation problem is performed by following what might be called a binocular approach: to gain perspective by examining two different economic spaces (the firm space⁵⁰ and the market space) and show that each contains analogous variables which can be taken to be identical with high probability.

In the firm space, value and surplus value are created in production; in the market space, value and surplus value are not only realised in exchange, but also – because of the purely probabilistic relations between (money) price and (labour) value – transferred among those who organise its production.

First consider the firm space: after arguing as described above for the gamma distribution of the rate of profit random variable R , Farjoun and Machover define a further random variable Z , the ‘rate of wage bill’, which is simply the annual gross wage bill of the firm divided by its capital. They appeal to the same heuristic argument to establish that Z is also gamma distributed with (at least) the same beta parameter as R . The random variable R/Z will then correspond to Marx’s notion of the rate of surplus value s/v , taken by him to be uniform. Farjoun and Machover appeal to a unique property of the gamma distribution to

⁴⁹ Note that the time span over which this so-called ergodic principle operates is long in relation to the micro events, but – in the physical systems considered here – is short by the standards of human observers.

⁵⁰ Or, as we prefer, the capital space.

show that if the random variable R/Z is single-valued then its only possible value is 1 (page 72).^{51,52}

Now consider the market space, the space of all market exchanges in the same time period as used to think about the firm space. Farjoun and Machover define further random variables

W : the money wage paid per unit of labour-time

: ‘specific price’, or the money price paid per unit of labour content, where the monetary unit is *defined* to be that which makes the average unit wage $E(W) = 1$, itself a ratio of two random variables, and .

: money paid in transaction i

: labour content of commodity traded in transaction i

Also, $(i) = V(i) + S(i)$, where

V : labour cost of commodity traded

S : profit realised on commodity traded

Farjoun and Machover show that $E(\) = E(W) + [S(i)/V(i)]$. By their choice of units $E(W) = 1$, and they argue that $S(i)/V(i)$ can be identified with R/Z (claimed also to be approximately 1, recall).

⁵¹ A point which Farjoun and Machover, surprisingly, do *not* mention, let alone stress, is that Marx always takes $s/v=1$; no doubt this makes the arithmetic of his examples simpler, but other values could give equal simplicity (0.5? 2? 100?), and perhaps Marx might have used one of these if he thought it more reasonable as a stylised fact.

⁵² See Appendix for details.

Thus, by considering the division of the price realised in each commodity transaction into payments to workers, and profits for capitalists, a market space analogue of the (statistical) equality of R and Z can be identified.

Although the first - $S(i)/V(i)$ - represents the division between capital and labour of the price of commodities traded, while the second - R/Z - is the division of value created, at least some of it in periods before that in which commodities are exchanged, Farjoun and Machover argue that 'if the economy is at or near dynamic equilibrium, the two ratios must be extremely close to each other, because the ratio between total profits and total wages cannot change rapidly' (1983:118), and indeed has already been shown to be very stable in the region of 1.

If a competitive market economy has a state of equilibrium, it must be a state in which a whole range of profit rates coexist; it must be a *dynamic* state, in the sense that the profit rate of each firm keeps changing all the time; it can only be a state of equilibrium in the sense that the *proportion* of capital (out of the total social capital) that yields any particular rate of profit remains approximately constant.

(Farjoun and Machover 1983:36)

Similarly they work in the market space (page 135) to show that the average rate of profit for the whole economy can be identified (once again, in a probabilistic way) with $E(R)$.

5.1.2 Farjoun and Machover on Marx

In relation to Marx's political economy, Farjoun and Machover's project can be described as one of both restoration and reconstruction. On the one hand, they see themselves as restoring the essence of Marx's thought, as expressed in *Capital* Volume I, and on the other as reconstructing his theory so as to rescue him from the *impasse* into which they believe he led himself in Volume III – that is, the transformation problem.

Our position is that Marx's political economy is fundamentally probabilistic; we shall suggest an alternative reading to Farjoun and Machover's which reveals Marx to have anticipated their perspective, albeit without the modern technical apparatus.

As Farjoun and Machover point out, there are two versions of the uniform profit rate hypothesis, a hard one and a soft one.

They begin (pages 14-15) with two quotations from Marx:

‘... there is no doubt, however, that in actual fact, ignoring inessential, accidental circumstances that cancel each other out, no ... variation in the average rate of profit exists between different branches of industry, and it could not exist without abolishing the entire system of capitalist production.’

(Marx, 1981: 252)

‘... capital withdraws from a sphere with a low rate of profit and wends its way to others that yield higher profit. This constant migration, the distribution of capital between the different spheres according to where the profit rate is rising and where it is falling, is what produces a relationship between supply and demand such that the average profit is the same in the different spheres, and values are therefore transformed into prices of production.

(Marx, 1981: 297)

These appear to support the hard version, formalised by Farjoun and Machover as the system of n input-output equations

$$P_0 = \mathbf{p}_i + R k$$

for each of n industries in economy, where P_0 is the price of a unit of output, \mathbf{p}_i is a vector of costs of inputs used up per unit of output, R is the uniform rate of profit and k represents capital per unit of output.

The soft version, in their words, says that ‘in a competitive economy, the values of [the rate of profit] for all different branches [of the economy] must be very close to each other, and for theoretical purposes can be taken as uniform across the whole economy’ (page 20) and since this doesn’t specify what is meant by a branch, if one takes large enough sectors giving, say, a dozen branches (and a long enough time period too) then this becomes realistic and plausible.

In fact, as they point out immediately after this, their quotations from Marx, read in context, support the soft and not the hard version of profit uniformity. We have also seen (Chapter Two) that they have no objections to this version. Hence it is difficult to accept their later statement (page 131) about the ‘irony’ of Marx’s acceptance of ‘the economists’ story of the alleged tendency of the rate of profit toward uniformity’. As we have seen they – and according to them, Marx – accept that this tendency is real.

Indeed, the real irony is twofold: not merely was Marx’s view of profit rate equalisation the one Farjoun and Machover prefer: that of an equilibrium distribution, but this view is part of a consistently probabilistic and statistical outlook throughout his career as a whole, and in *Capital* in particular. Indeed, Farjoun and Machover themselves recognise this elsewhere when they quote Marx’s claim that ‘the true law’ of economics is ‘*chance*’ (Farjoun and Machover, 1985). We now examine Marx’s views in more detail.

5.1.3 Marx’s probabilistic political economy

We claim that Farjoun and Machover’s approach is faithful to Marx in so far as the latter’s approach was clearly and consciously statistical from first to last. Here we indicate some evidence for this, beginning with that which they themselves bring forward. The claim that ‘the true law of economics is chance’ is made very early in Marx’s career, in the ‘Notes on James Mill’ written in 1844:

[I]n his demonstration that the cost of production is the sole factor in the determination of value Mill succumbs to the error ... of defining an *abstract law* without mentioning the fluctuations or the continual suspension through which it comes into being. If *e.g.* it is an

invariable law that in the last analysis - or rather in the sporadic (accidental) coincidence of supply and demand – the cost of production determines price (value), then it is no less an *invariable law* that these relations do not obtain, *i.e.* that value and the cost of production do not stand in any necessary relation. Indeed, supply and demand only ever coincide momentarily thanks to a previous fluctuation in supply and demand, to the disparity between the cost of production and the exchange value. This is the *real* movement, then, and the above-mentioned law is no more than an abstract, contingent as one-sided movement in it. Yet recent economists dismiss it as accident, as inessential. Why? Because if the economists were to attempt to fix this movement in the sharp and precise terms to which they reduce the whole of economics this would produce the following basic formula: laws in economics are determined by their opposite, lawlessness. The true law of economics is *chance*, and we learned people arbitrarily seize on a few moments and establish them as laws.’ (Marx, 1975a: 259–260; emphases in original)

Farjoun and Machover (1985) quote only the final two sentences of this,⁵³ in support of the mild claim that ‘[e]conomists and economic philosophers have often pointed out the essentially indeterminate and statistical nature of economic categories such as price and rate of profit’. Taken in isolation this might be understood as little more than the claim that good-sized confidence intervals should be installed around the claimed values of any economic data.

But it can be argued that Marx is not talking about the accuracy of data, but about the nature of the process which causes it to have one value rather than another. In fact Marx’s interest in the role of chance is manifest in his very first serious intellectual production, his doctoral dissertation (Marx, 1975b). In this work, Epicurus’ physics is praised over that of Democritus on the specific ground that Epicurus introduces chance as a way of making room for human free will (McLellan, 1980).

⁵³ In slightly different wording, because they cite the *Collected Works*.

Moreover, the importance to Marx of a probabilistic methodology of economics can be seen from the fact that the ideas of the notes on Mill are repeated in his first published comments on political economy. In the pamphlet *Wage labour and capital* – originally written as a lecture to the German Workingmen's Club in Brussels in 1847 – he summarises his thought thus: 'the total movement of this disorder is its order' (Marx, 1952). Still clearer is the formula in *The poverty of philosophy* (Chapter One, section 2), written in the winter of 1846-47 against Proudhon:

If M. Proudhon admits that the value of products is determined by labour time, he should equally admit that it is the fluctuating movement alone that in a society founded on individual exchanges makes labour the measure of value. *There is no ready-made constituted 'proportional relation', but only a constituting movement.*

(Marx, n.d.: 71; emphases added)

With this in mind we turn to Marx's comments in *Capital* Volume III:

Competition distributes the social capital between the various spheres of production in such a way that the prices of production in each of these spheres are formed after the model of the prices of production in the spheres of mean composition, *i.e.* $k + kp'$ (cost price plus the product of the average rate of profit and the cost price) ... the rate of profit is thus the same in all spheres of production, because it is adjusted to that of these average spheres, where the average composition of capital prevails.

(Marx, 1981: 273)

On the face of it this passage – so far – appears to support the 'hard' version of profit uniformity discussed by Farjoun and Machover. But Marx continues:

... Between those spheres that approximate more or less to the social average, there is again a tendency to equalization, which seeks the 'ideal' mean position, *i.e.* a mean position which does not exist in reality. In other words, it tends to shape itself around this ideal as a norm.

(Marx, 1981: 273)

Two points: first, Marx says that the ‘ideal’ mean position *does not exist*: this is an echo of the distinction drawn by Quetelet between a real average (*moyenne*) and an arithmetic average (*moyenne arithmétique*) (Mosselmans, 2005). The former denotes, for example, finding the position of a star as an average of several observations (there really is just one star there, and at some particular place rather than another). The latter denotes such activities as finding the height of the ‘average man’, as Quetelet dubbed him, by averaging the heights of a number of different men; but there may be no actual man whose height is equal to the average.⁵⁴ Marx read Quetelet (the International Institute for Social History in Amsterdam has a notebook containing his reading notes on Quetelet’s *Treatise* (International Institute for Social History, 2006)) and cited him elsewhere, as we will see below. In this light it is hard to read Marx as asserting an actual equalisation, even within those spheres that ‘approximate more or less to the social average’.

Second, Marx describes the tendency to profit rate equalisation as a tendency to *shape itself around* this ideal [mean] as a norm (our emphasis); profit rate equalisation *is* the ‘shape around the ideal’ towards which the tendency is directed – in other words, it is the formation of a profit-rate distribution (‘shape itself around’ the mean).

This might be thought a fanciful interpretation were it not that a few pages later, when Marx discusses intra-industry variations in productivity (the difference between the individual value of a commodity and its social value), he not only describes a probability density but also discusses how variations in its shape – symmetric or not, light- or heavy-tailed – will affect the relation of the mean to the whole; he even considers the effect of censoring some of the data (1981: 283–284).

⁵⁴ The status of the ‘average man’ is controversial: Quetelet thought that the individual who most nearly attained to the average in every faculty would represent the genius of the age: ‘A man can have no real influence on masses – he cannot comprehend them and put them in action – except in proportion as he is infused with the spirit which animates them, and shares their passions, sentiments and necessities, and finally sympathises completely with them. It is in this manner that he is a great man, a great poet, a great artist. It is because he is the best representative of his age, that he is proclaimed to be the greatest genius.’ (Quetelet, 1842 [1969]).

In thus considering distributions that might be either non-symmetric or heavy-tailed or both Marx was well ahead of the ideas of many professional statisticians of his day. The Gaussian distribution acquired its misleading alternative designation as ‘normal’ precisely because many thought it was the distribution to be expected in most social and biometric data; it is of course symmetrical, and in the usual Pearsonian system of moments the description of a distribution as having ‘light’ or ‘heavy’ tails is relative to the Gaussian kurtosis. Although the gamma distribution was first derived by Laplace as far back as 1836 (Johnson *et al.*, 1994: 343), it became better known after it was revived by Pearson as part of his system of distributions (in which it is Type III) in 1893. Even then there was considerable resistance to idea that asymmetric distributions might be a useful tool (Stigler, Stephen M., 1986: 333–341).

Finally we come to Marx’s concept of social labour. This is the only point where Marx refers to Quetelet in connection with definitely economic ideas: in the first footnote to Chapter 13 of *Capital* Volume 1 Marx cites Quetelet (and Edmund Burke) in support of his explanation of his concept of average social labour:

Edmund Burke, that famous sophist and sycophant, goes so far as to make the following assertion, based on his practical observations as a farmer: that ‘in so small a platoon’ as that of five farm labourers, all individual differences in the labour vanish ... For example, let the working-day of each individual be 12 hours. ... From the point of view ... of the capitalist who employs these 12 men, the working-day is that of the whole dozen. ... *But if the 12 men are employed in six pairs, by six different ‘small masters’, it will be entirely a matter of chance whether each of these masters produces the same value, and consequently whether he secure the general rate of surplus-value. ... The inequalities would cancel out for the society as a whole, but not for the individual masters.*

(Marx, 1976: 440–441)

For Marx, and the capitalist, the product of the working day is the total labour of the workers employed; this is a random variable found by summing the random variables

constituted by the individual labours of each worker, which are social labour only in so far as they are employed by capitalists employing other social labours.

Here too we see evidence that Marx's probabilistic approach put him, implicitly, in the vanguard of contemporary statistical work; his description of the output of the collective working day as a sum of random variables is an informal version of the central limit theorem (CLT); although this was moderately advanced for the mid-nineteenth century it was hardly novel, a version of the CLT having first been described by de Moivre in the seventeenth century and revived by Laplace at the beginning of the nineteenth (Stigler, Stephen M., 1986: 136). The ordinary version of the CLT covers only sums of distributions with finite variance, that is, with light tails. We have seen that Marx contemplated heavy-tailed distributions; moreover the CLT points to the normal distribution as that of the summed variables, which is symmetrical, whereas the distribution of the value output of a 'gang' of workers is surely bounded below by zero. However, the generalised central limit theorem, which relaxes the restriction to finite variance and allows asymmetric, heavy-tailed distributions as the outcome, was not developed until the 1920s (see also Chapter Seven).

5.1.4 Testing Farjoun and Machover

In this section we discuss various considerations relating to practical tests of Farjoun and Machover's system, and in particular their hypothesis that the rate of profit should have a gamma distribution.

We begin by noting that their full hypothesis is a complex one about the rate of profit R and their rate of wage bill variable Z .⁵⁵ As discussed in section 5.2.1 above, they argue that the random variable R/Z , the analogue of the rate of surplus value, should be an almost-

⁵⁵ In fact, a complete test of their system would also involve their propositions about the market space. But such a test would require information about the prices paid in individual transactions for many commodities. In principle one might do this, since exactly such data is now collected and stored electronically by both manufacturers and distributors (so-called 'data warehousing').

degenerate one with value = 1. In fact, the full argument on this point also involves the hypotheses that R and Z can both be expected to be gamma distributed with a common scale parameter λ , and that if their shape parameters are α_R, α_Z , then the random variable $R + Z$ should have a gamma distribution with parameters $(\alpha_R + \alpha_Z, \lambda)$.

It is clear from their discussion that it is only the wages of productive workers which should be included in Z . Unfortunately our FAME database does not allow us to recover this information directly. It is included in the cost-of-sales variable $COST$ (which is why we are able to calculate Gillman's Marxian measures of the profit rate), but the FAME variable $REMU$ covers wage-related payments to all employees, not just those who are productive, in the Marxist sense. For this reason we confine ourselves to testing the rate of profit aspect of Farjoun and Machover's work.

Given that we are going to test a variety of profit rate measures we have to consider a range of possible outcomes. We argued in Chapter Two that Farjoun and Machover's discussion of what they understand by 'rate of profit' suggested that their conception was most closely matched by the measure we have labelled Gillman 4 (or, less confidently, Gillman 3). However, more than one measure might turn out to have a gamma distribution, and these might or might not include Gillman 3 or 4.

There is also the question of which version of the gamma distribution we consider. Farjoun and Machover (1983) consider only those versions of the gamma with two parameters (shape and scale), but as shown in the Appendix the general form of the gamma includes a location (threshold) parameter and a secondary shape parameter. Farjoun and Machover explicitly discount consideration of gamma distributions with a negative threshold, on the grounds that loss-making firms are not a significant feature of the economy (1983:66-67).⁵⁶ This is, firstly, because they say that firms which makes losses are not only few in number, but generally of small size, and thus account for only a small

⁵⁶ Implicitly they also rule out positive threshold parameters.

proportion of the total social capital. Second, their profit rate is calculated before distribution of extra-firm claims such as interest, and few firms will make losses on this basis.

This might be thought excessively restrictive; it can be expected that there will always be firms which make losses, either because they are start-ups, or because they are in danger of failing. Either condition can be viewed as a normal part of the competitive process. If the threshold is treated as a parameter to be estimated, then it has an obvious economic interpretation as the maximum rate of loss compatible with survival, where the fact of survival is determined by social capital's assessment of the stewardship of the firm's managers (see our discussion of Bryer (1994) in Chapter Two).

Although Farjoun and Machover discount the idea of the profit rate distribution having a negative lower bound they do not explicitly rule out positive values for the location parameter (although their own small-scale test does assume location at zero).

Because the four-parameter gamma distribution includes a number of other distributions as special or limiting cases (the latter including the log normal), one might consider both the Farjoun and Machover and Gibrat hypotheses to be special cases of a more general hypothesis yet to be formulated.⁵⁷ According to Lienhard and Meyer (1967) the second shape parameter is related to subsidiary details in a number of processes known to generate such distributions as limiting outcomes over time.⁵⁸

We regard the question of the location of the distribution (and also its scale to be fundamentally secondary to the question of its shape. If the distribution of a profit rate

⁵⁷ McDonald introduces the generalised gamma in connection with income distribution. A much earlier proposal in this connection is that of Amoroso (1925), so Gibrat (1931) was arguably out-of-date on publication.

⁵⁸ Some other uses of this distribution in economics include Creedy *et al.* (1994), as a model for the distribution of earnings in the labour market, and Creedy and Martin (1994), as a model for the distribution of market prices in general.

measure can be shown to have a shape which *cannot* be described by any member of the generalised gamma family, appropriately scaled and shifted, we will interpret this as *strong* disconfirmation of Farjoun and Machover's hypothesis; if the distribution *can* be shown to be describable by *some* member of the generalised gamma, but not the three-parameter (shape-scale-location) version, we will regard it as *weak* disconfirmation. Table 5.1 summarises the possibilities.

Table 5.1: testing Farjoun and Machover (1983)

Gillman 3/4 is gamma (n° of parameters)	Alternative measure is gamma	Farjoun and Machover confirmed?	Comment
Yes (two)	Not discussed	Yes	Small-scale test in Farjoun and Machover (1983)
Yes (three)	Yes	'Strong' confirmation	Ideal gas metaphor has profound implications for theory of capitalist competition (1)
Yes (three)	No	'Neutral' confirmation	Ideal gas metaphor is useful and profit rate concept supported (2)
No	Yes	'Weak' confirmation	Ideal gas metaphor is useful, but profit rate concept must be re-thought (3)
Yes (four)	Yes/no	'Weak' disconfirmation	Strict interpretation of Farjoun and Machover disconfirmed; their hypothesis a special case of a wider model (4)
No	No	'Strong' disconfirmation	Ideal gas metaphor not useful, but no particular implications for 'econophysics' approach in general (5) No particular implications for 'marxist' concept of the profit rate (6)

What are the implications of these various possible outcomes?

(1) 'Strong' confirmation: if not only Gillman 3/4 but also one or more conceptually very different measures turn out to have gamma distributions then the ideal gas metaphor would seem have quite profound implications for our understanding of capitalist competition and the way in which claims on surplus value arise.

(2) 'Neutral' confirmation: if only Gillman 4 (and/or Gillman 3, possibly) has a gamma distribution, then Farjoun and Machover are clearly vindicated. Subsequent work to explain the distributions displayed by other measures would clearly be of interest.

(3) ‘Weak’ confirmation: if one or more profit rate measures had gamma distributions, *but not* Gillman 3/4, this would support the ideal gas metaphor as a way of thinking about capitalist competition, but would also suggest that the appropriate definition of the profit rate needed to be rethought. Farjoun and Machover’s hypothesis about the profit rate distribution does not depend on any formal model of the distribution generating process, and certainly not on one relying on particular properties of the rate of profit they discuss.

(4) ‘Weak’ disconfirmation, discussed above; the strict interpretation of Farjoun and Machover cannot be confirmed, but a broader interpretation can be sustained.

(5) ‘Strong’ disconfirmation would not necessarily discredit the general approach of the ‘econophysics’ movement discussed in Chapter One (the application of the statistical mechanics paradigm to economics). But the distribution(s) actually found would clearly have to be accounted for in a different way to that advanced by Farjoun and Machover.

(6) Finally, strong disconfirmation would seem to have no particular consequences for one’s view of which profit rate measure was appropriate for empirical work in marxist economics.

5.2 Distributions of rates of return

In this section we apply our L -moment ratio analysis to profit rate distributions, investigate the properties of the sampling process by which they are derived, and show that for all the profit rate measures investigated both the size-weighted and unweighted versions have distributions with extended tails displaying power law characteristics. We conclude with a discussion of the further lines of research suggested by these results.

5.2.1 L -moment ratio analysis of profit rate distributions

Here we apply the methods developed in Chapter Four to investigate the distributional models of different profit rate measures; we use randomly-sized random samples to estimate the L -skewness γ_3 and L -kurtosis γ_4 for the data from each of the five years 1991–5, plot the resulting clouds of points and assess their relationship to the loci of the

distributions given by Hosking and Wallis (1997). As in previous chapters we group the results in four sections: the marxian measures defined by Gillman (1956), the same author's 'capitalist' measures, the four standard accounting ratios, and the eight measures discussed in Glick (1985).

To assist comparison we begin by showing the clouds of points in a common sub-set of the total L -skewness, L -kurtosis space, namely $-0.25 \leq s_3 \leq 1$ and $0 \leq s_4 \leq 1$ (Figures 5.1 to 5.4); for reference purposes we include the loci of distributions previously shown in Figure 4.1, as provided by Hosking and Wallis (1997).⁵⁹ The points representing each year are identified by the appropriate final digit.

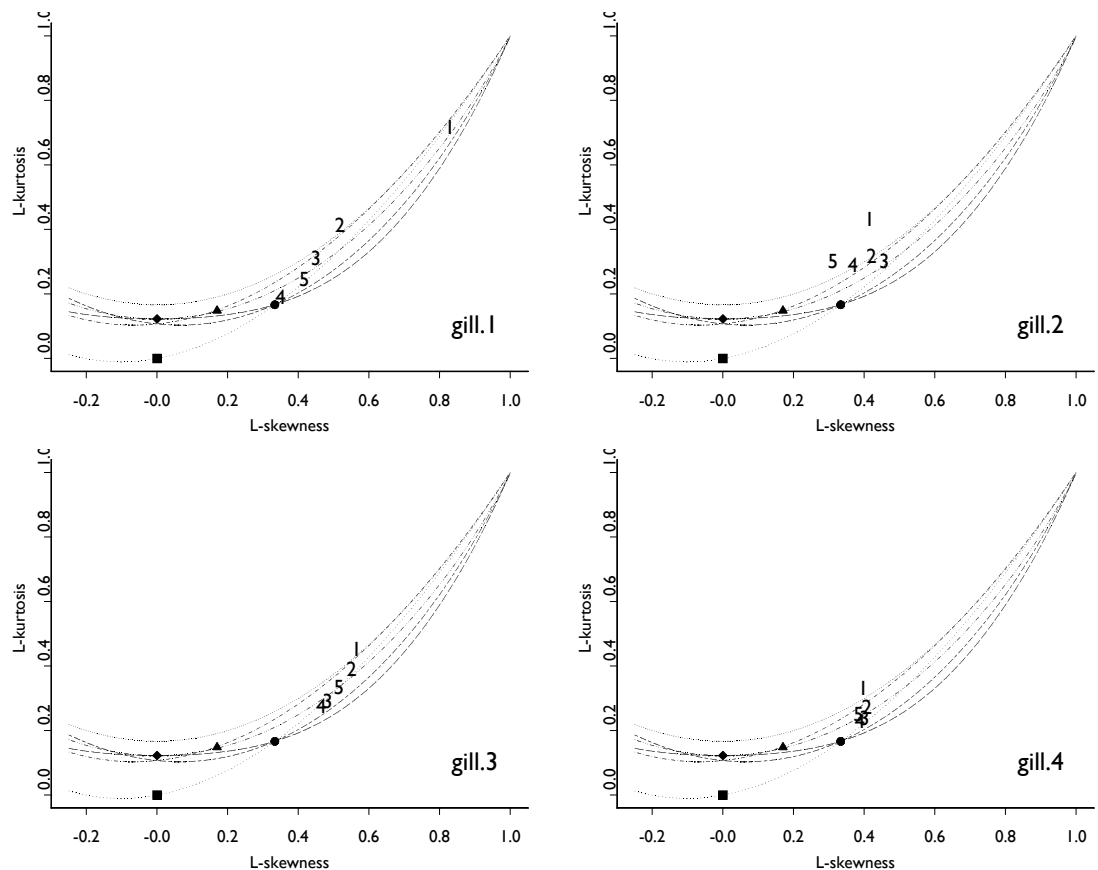


Figure 5.1: L -skewness and L -kurtosis of Gillman marxian measures

⁵⁹ To recap, these are the uniform (identified by ■), the Gaussian (◆), the exponential (●), the Gumbel (▲), and the linear loci of six three-parameter distributions: from top to bottom at the $s_3 = -0.25$ ordinate they are the generalised logistic, the generalised extreme value, the three-parameter log normal, the three-parameter gamma, the Weibull, and the generalised Pareto.

All four of Gillman's capitalist measures fall within the broad band encompassing Hosking and Wallis's loci (Figure 5.1). However, the point clouds associated with each measure – apart perhaps from Gillman 4 – are too widely-scattered to be readily associated with the locus of any particular one of the distributions shown. Gillman 4 is the measure we identify as the one closest to the profit-rate concept adopted by Farjoun and Machover (1983), so its relatively-tight clustering appears encouraging. In the next section we will investigate this more closely.

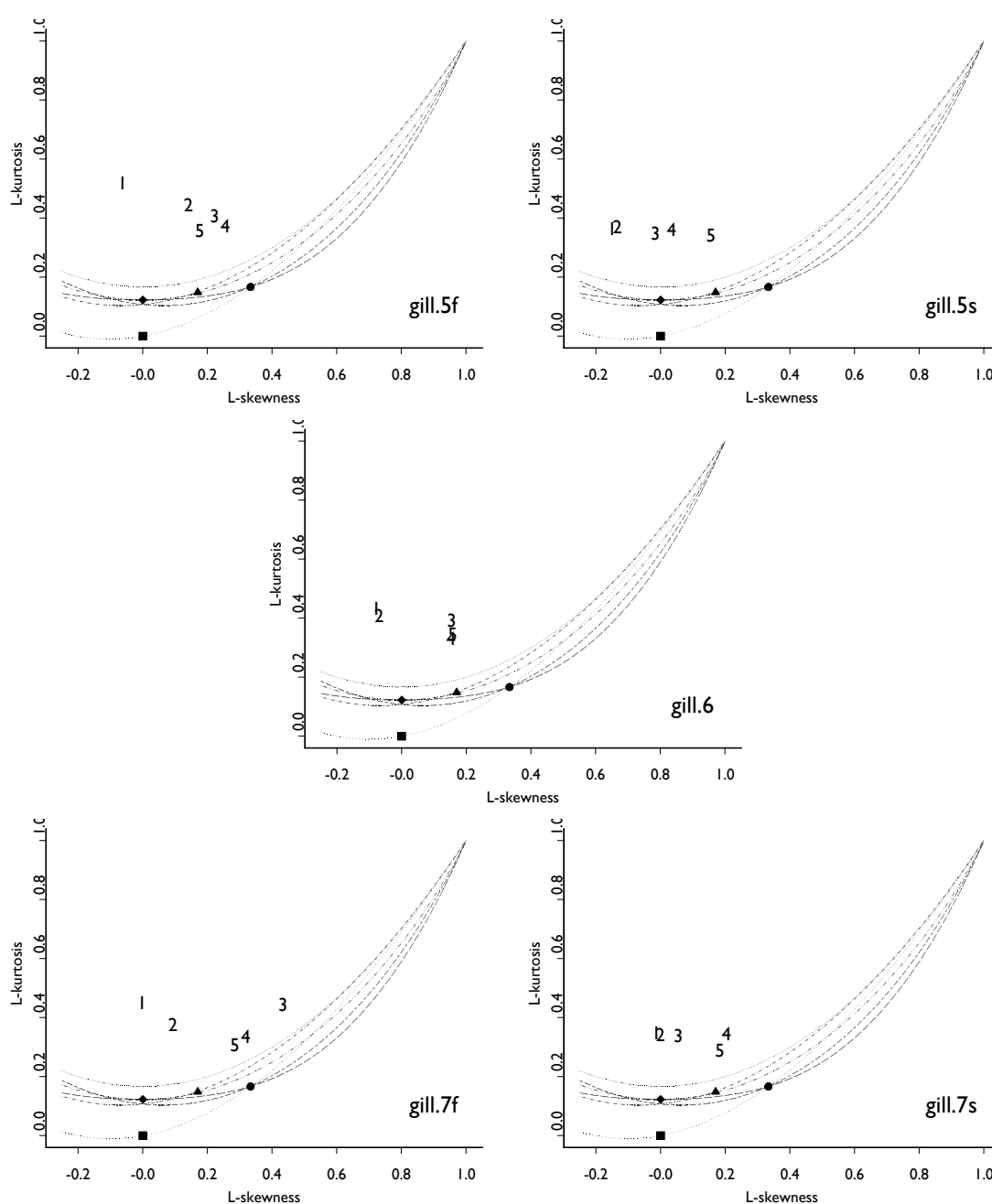


Figure 5.2: L-skewness and L-kurtosis of Gillman capitalist measures:

In contrast, Gillman's 'capitalist' measures all display greater L -kurtosis, relative to L -skewness, than that attained by any of the Hosking and Wallis distributions (Figure 5.2). Moreover, their kurtosis is, relatively, little-changed from year to year, unlike their skewness, which displays a much wider range than for any of the marxian ratios apart from Gillman 1. A further notable contrast is that they all have years in which the estimated L -skewness is either close to zero or actually negative.

An intriguing feature of Gillman 5s is that the skewness is date-related, being at its smallest in 1991, then increasing monotonically through 1995; others of these measures have a partially-similar ordering.

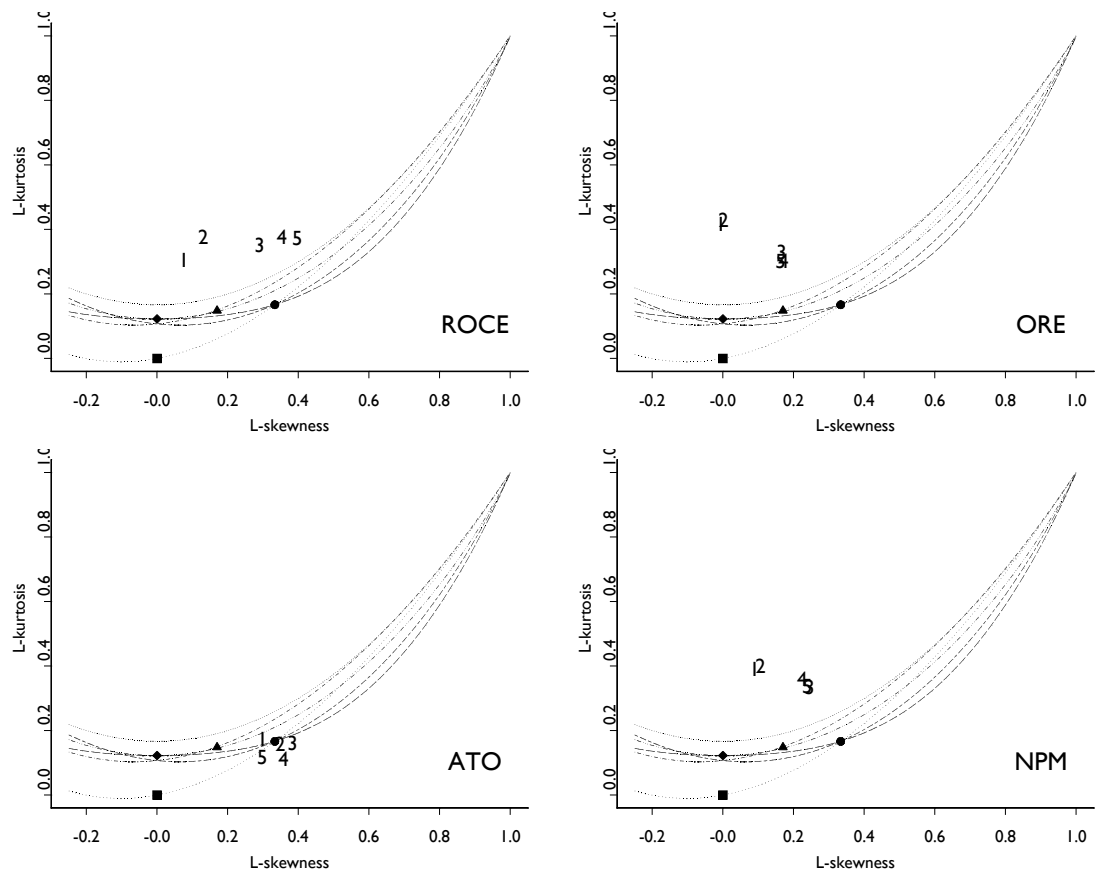


Figure 5.3: L -skewness and L -kurtosis of accounting ratios

Turning to the accounting ratios we find similar results to those for Gillman's 'capitalist' ratios (Figure 5.3): high kurtosis which is relatively unvarying in comparison to skewness, which again shows signs of being date-related.

The sole exception is the asset turnover ratio (ATO). We know from work in the previous two chapters that this measure is unusual: in unweighted form it is clearly bimodal (Figure 3.5), while the weighted version displays a tail which also hints at the existence of a second mode (Figure 4.5). However, if one neglects this tail the histogram does suggest an exponential form, which is consonant with the clustering around the locus of the exponential distribution (•) seen in Figure 5.3.

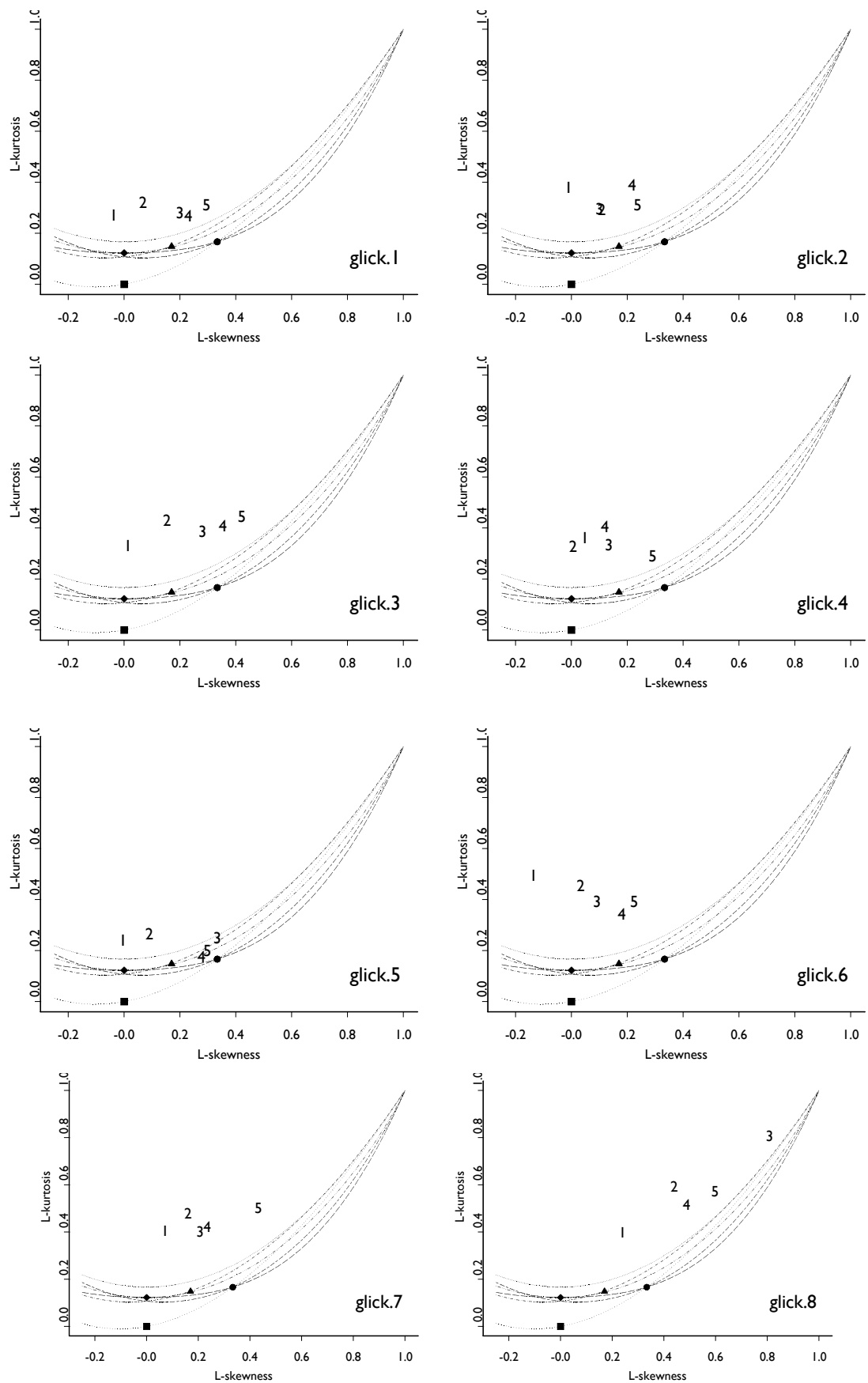


Figure 5.4: L-skewness and L-kurtosis of Glick measures

Finally, the majority of the Glick measures show results similar to those of the accounting ratios, unsurprisingly given their close conceptual relationships. The exceptions are Glick 5 (relatively low levels of L -kurtosis) and Glick 8 (wide range of L -kurtosis, which is in the main closer to the level corresponding to its L -skewness in the Hosking and Wallis distributions). However, even these display the tendency for skewness to be time-related. We will return to this in section 5.4.3 below.

We have seen that of all the 21 profit rate measures tested no less than 19 seem to fall outside the loci of those distributions we can identify using Hosking and Wallis's L -moment ratio system. In particular, they do not seem able to be associated with any case of the four-parameter gamma distribution. In the scheme set out in Table 5.1 above we thus have either row 2 ('neutral' confirmation of Farjoun and Machover, with either Gillman 3 or 4 found to be gamma), or row 4, 'weak' disconfirmation (with these two Gillman profit rate measures not gamma distributed). However, we suggested that disconfirmation could also come in strong or weak forms, with the latter consisting of showing that although Gillman 3/4 could not be modelled by a three-parameter gamma distribution, one of the special cases of the four-parameter gamma would provide a model.

As seen in Figure 5.1, Gillman 4 does appear to be a candidate for association with a particular distribution using Hosking and Wallis's L -moment ratio system (so does ATO, but this is unlikely by reason of other information about its form). Moreover four of the distributions for which loci are shown in Figure 5.1 are special cases of the four-parameter gamma (the exceptions are the generalised versions of the logistic and Pareto). Hence the locus of the four-parameter gamma is the region bounded by the envelope of these four distributions.

Thus if Gillman 4, the definition most strongly associated with Farjoun and Machover's hypothesis of a gamma model, can be shown to have either a generalised logistic or generalised Pareto distribution we have strong disconfirmation, and if one of the others, weak disconfirmation. We therefore investigate it in more detail.

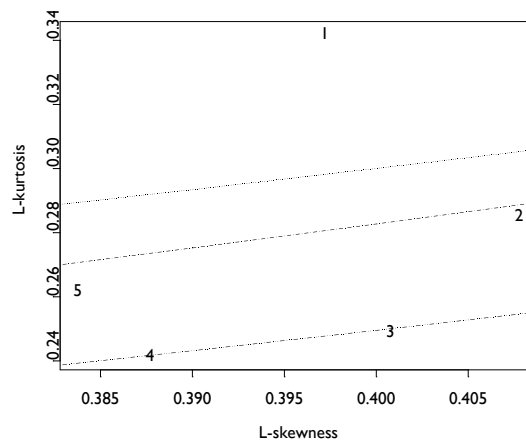


Figure 5.5: L -skewness and L -kurtosis of Gillman 4 in closer focus

Figure 5.5 focuses on the region of L -skewness, L -kurtosis space containing the observations on Gillman 4. Four years' observations fall in a region bounded above by the locus of the generalised extreme value distribution and below by the locus of the log-normal distribution. Since these are both special cases of the four-parameter gamma this is evidence of weak, as opposed to strong, disconfirmation of Farjoun and Machover (but this is equivalent to confirmation of a broad interpretation of their hypothesis).

However, the absolute variation in L -skewness is extremely small compared to that in L -kurtosis (masked here by the proportions of Figure 5.5, but more evident in Figure 5.1). Also, the fifth observation, that for 1991, has L -kurtosis considerably higher than for the other years, being well above the locus of the generalised logistic distribution.

Clearly we would like to confirm or reject these hypotheses more definitely. There is also the anomalous observation for 1991 to take into account.

5.2.2 Sampling properties of RS2 estimation

As a first step we investigate the sampling properties of our RS2 estimation method in the case of the 1995 data for Gillman 4 (Figure 5.6). Although they will not be reported here, investigations similar to those described below have been made for the other profit rate measures in our study; their results are qualitatively similar and suggest similar conclusions to those that will be drawn below.

It will be recalled that the RS2 sampling procedure involves taking 100 samples from our data, samples in which the probability of any company's profit rate value being included is proportional to the company's size relative to the largest company, as measured by the capital definition involved in the relevant profit rate measure.

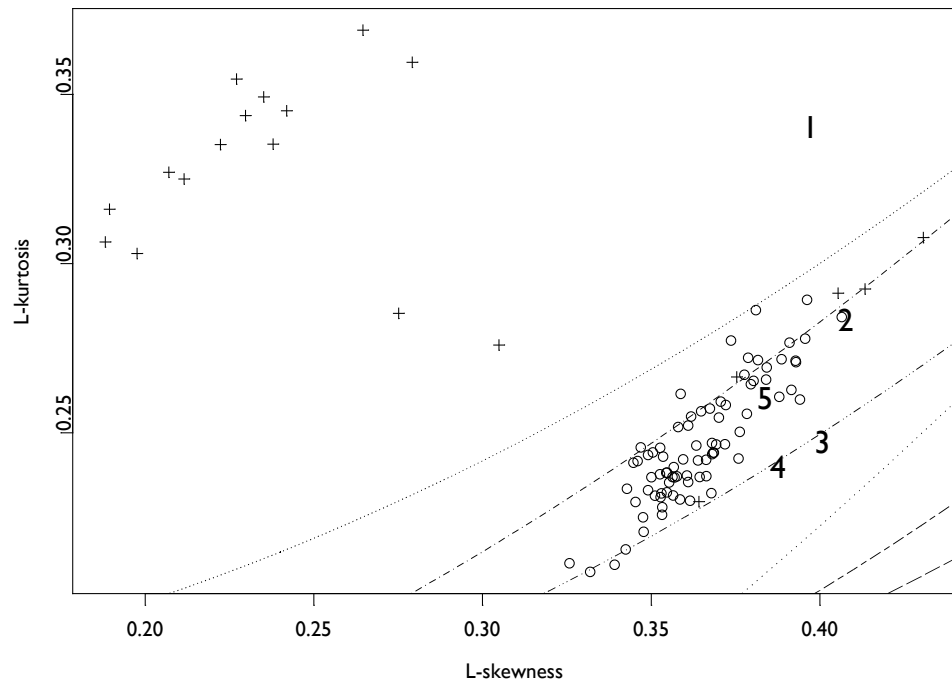


Figure 5.6: Gillman 4, 1995; L -skewness and L -kurtosis of RS2 samples

We begin by plotting the L -skewness and L -kurtosis of the 100 RS2 samples from 1995. Figure 5.6 shows the majority of samples as open circles; crosses indicate samples identified as discordant by reason of their distance from the main body of samples, using Hosking and Wallis's suggested test (1997:45ff). The estimated L -skewness and L -kurtosis for each year of our data are plotted by the appropriate final digit.

Several features are notable. First, the samples form two distinct clouds, a main one to the south-east of the plot and a subsidiary cloud of some dozen points to the north-west. The reason for this is unknown but we believe that it is a random outcome, as it did not appear in a re-run of the sampling procedure. More extensive investigation of this was deferred on the grounds of the heavy computational load.

Second, consider the ranges of L -skewness and L -kurtosis exhibited by the samples: both more than cover the range of annual variation in estimated values, even when the anomalous result for 1991 is included.⁶⁰ The difference is especially pronounced in the case of L -skewness.

Third, if we consider only the main cloud of samples, the distributions of both L -moment ratios have pronounced skewness, with upper tails longer than the lower.

Fourth, the tails produce an impression of strong correlation of the sample L -moment ratios.

Fifth, the estimated L -moment ratios for 1995 lie outside the core of the main cloud of points (which, recall, represent samples from the 1995 data).

Taking the last point, if one assumed that this divergence between the samples and the estimates to which they give rise results from the skewness of the sample L -moment ratios, that this result was also a feature of other years' data, and that some appropriate correction would result in the L -moment ratio estimates being translated southwards, then the estimates for 1992–5 might well lie athwart the locus of the log-normal distribution (confirming Gibrat's hypothesis), and even the estimates for 1991 would look less anomalous (a point which would hold even more strongly if one assumed a translation proportional to the starting values).

To see how far such a corrective translation might be justified, and also to see the reason for the skewness of the sample L -moment ratios, we now look directly at the sample L -moments, as opposed to the L -moment ratios.

⁶⁰ The presence of a sub-group of samples with relatively low skewness and high kurtosis should also be noted, but investigating this was judged to take us too far from the objectives of our present study.

The fact that the estimated skewness and kurtosis both lie outside the main mass of sample values is not (directly) a result of the skewed distribution of the samples. To see this, recall that we estimate L -skewness and L -kurtosis *not* by averaging the sample values of these ratios, but by averaging the sample moments and taking the ratio of the appropriate results.

This is illustrated in Figure 5.7, where we plot cross-sections of the four-dimensional space containing the first four L -moments: μ_1 (the mean) and μ_2 , μ_3 and μ_4 .

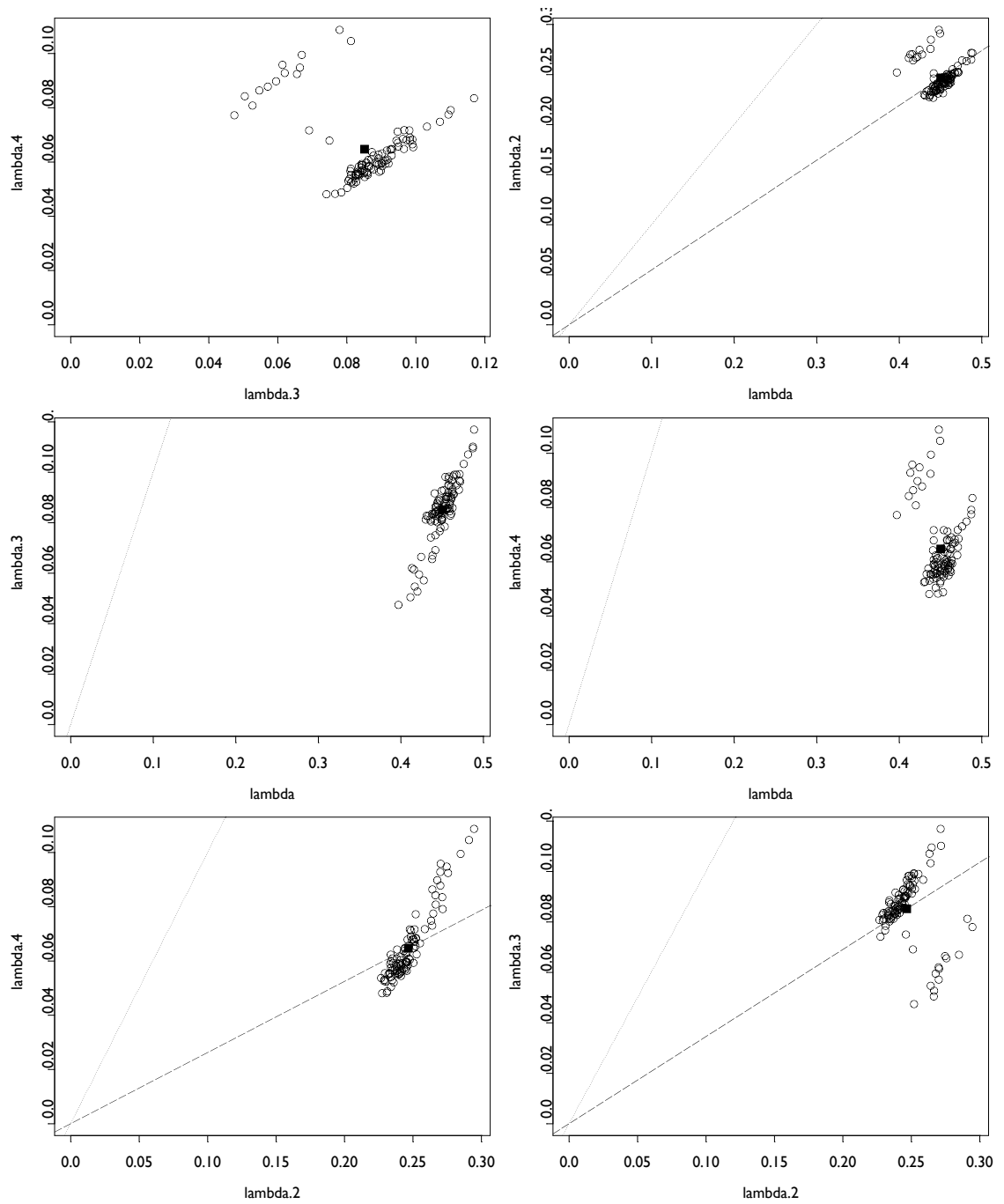


Figure 5.7: Gillman 4, 1995; L -moment space of RS2 samples

Here the solid squares identify pairs of estimated L -moments, and dotted lines indicate the 45° degree line. In the λ_2 , λ_3/λ_2 and λ_4/λ_2 panels, dashed lines indicate a ray from the origin through the estimates of the population L -moments; the relevant ratio is thus measured by the gradient of the ray.

The first panel (top row left) is the λ_4/λ_3 plane. This L -moment cross-section bears a strong resemblance to its L -moment ratio counterpart. The reasons for this can be seen

from examination of the other panels. First, while in each of these we see a similar pattern to that shown by the L -moment ratios – strong correlation between each pair of moments – the apparent degree of association is weakest in the μ_3/μ_2 plane. Second, in this case the projection of the axis of the four-dimensional point-cloud lies approximately along the 45° line, whereas in the other planes the axis of the cloud is displaced (while still remaining approximately parallel to the 45° line). Third, note the very strong relationship between the third and fourth moments, on the one hand, and the mean on the other (second row of Figure 5.7). Small proportional changes in the mean of the sample imply much larger changes in the third and fourth moments, but little change in the second moment (or in the coefficient of L -variation). Since the L -skewness and L -kurtosis are the ratios of the third and fourth moments, respectively, to the second, the implication is that these ratios are primarily sensitive to differences in the means of the higher moments of the samples.

The third row of Figure 5.7 shows the μ_3/μ_2 and μ_4/μ_2 cross-sections. Here we note the skewness shown by all three higher moments (but to a lesser degree by the mean).⁶¹ The skewness of the sample moments obviously biases the estimates of each moment, which in turn biases the estimates of the ratios because of the high correlation among the moments. In principle, one might try to mitigate this by omitting samples identified as discordant by Hosking and Wallis's test. Doing so would tend to change the estimated value of each moment – but because the ratios are estimated by the slopes of the rays, it is also clear that this would be unlikely to make a significant difference to the estimates of the ratios (essentially because of the shallow angle between the axis of the four-dimensional cloud of points and the 45° degree line). In the case of Gillman 4 it appears that the L -skewness might actually *increase* slightly, rather than decrease: but because of the positive slope of the lognormal locus this would tend to improve the conformity to this locus, for any given decrease in L -kurtosis.

⁶¹ We also again note the off-axis samples, which themselves show signs of mutual correlation.

We conclude that the relative lack of success of our RS2 and L -moments procedure in clearly identifying distributional models is not due to problems with the procedure.

Instead, recall that our RS2 sampling scheme is designed to attenuate the influence of very small companies, which we know are associated with very wide ranges of profit rate, and which we suspect are often not genuine capitalist entities. However, it is not designed to exclude them entirely, since we have no specific information on which to base the exclusion of any particular company.⁶² Thus our sampling procedure will occasionally create samples which include a disproportionately high number of extreme values which give an upward bias to the sample's L -moments.

As we have shown, this does not in practice impart undue bias to the skewness and kurtosis estimates. However, the prevalence of samples containing extreme values is shown by the pronounced skewness of the sample of the higher L -moments. This directs renewed attention to the tails of the distributions, both weighted and unweighted. The relatively high prevalence of extreme values among the moments of the samples points to the fact that even the size-weighted version of this profit rate measure has a heavy upper tail.

By hypothesis, each year's profit rate observations are drawn from a common distribution (neglecting variations in parameters); *a fortiori* the samples from a given year's data are drawn from a common distribution. The latter evidently have wide variation in skewness and kurtosis, which we have shown to be linked to variation in the mean, itself subject to quite wide fluctuation due to the generation of relatively many samples with many extreme values.

We therefore conjecture that the observed large annual fluctuations in kurtosis for Gillman 4 are principally due to fluctuations occurring in the tails of the distribution, not in

⁶² If we did, we could exclude it directly.

the main portion. (By extension, we conjecture that variation in the tails drives the annual variations in skewness and kurtosis observed for other profit rate measures.)

One could more confidently assert a four-parameter gamma model, bounded by the log normal and generalised extreme value distributions and thus weakly disconfirming Farjoun and Machover but confirming a more general version of their hypothesis, if one had some justification for omitting the tails from the analysis. We do not know of any such justification on *a priori* grounds, but a more detailed empirical investigation may suggest possible lines of approach. Thus in the next section we examine the tail structure of all our profit rate measures more closely.

5.3 Zipf plot analysis of tails of profit rate distributions

We now investigate the tails of the distributions using Zipf plots, in which the log of the order statistics is plotted on the horizontal axis and the log of their rank on the vertical.⁶³ Zipf plots are so-named for their ability to illustrate Zipf's Law (Zipf, 1932), a proposal about word frequency in texts; the Zipf distribution is closely related to the Pareto distribution. The plots were, apparently, introduced into economics by Stanley *et al.* (1995) in the analysis of firm-size distributions.

In a Zipf plot distributions from the exponential family (of which the gamma distribution is a member) have upper tails curved at all points. Power-law distributions, such as the Pareto, have straight tails. This is illustrated in Figure 5.8, which compares (a) the standard Gaussian, (b) the standard exponential, (c) a gamma distribution with shape = 2 and scale = 1, (d) a Pareto distribution with location = 10^{-2} and scale = 1.75, and (g) the standard lognormal distribution. For each distribution we plot a random sample with $n = 10^5$.

⁶³ In the interests of speeding both plotting and printing, we modify the full Zipf plot as follows: we show the 100 largest values, and then 1,000 smaller values evenly-spaced as to rank; further, since our empirical data varies in length we shift the plots of the shorter series upwards to give the plot of each series a common starting point as to rank.

Since we plot the log of the values we are of course only considering each distribution's support on the positive real line. The existence of cases where the L -skewness is negative suggests that a fuller analysis should also take into account the left-hand tails (see Tables 3.4, 3.5 and 4.2, and Figures 3.7 to 3.11 and 4.10).

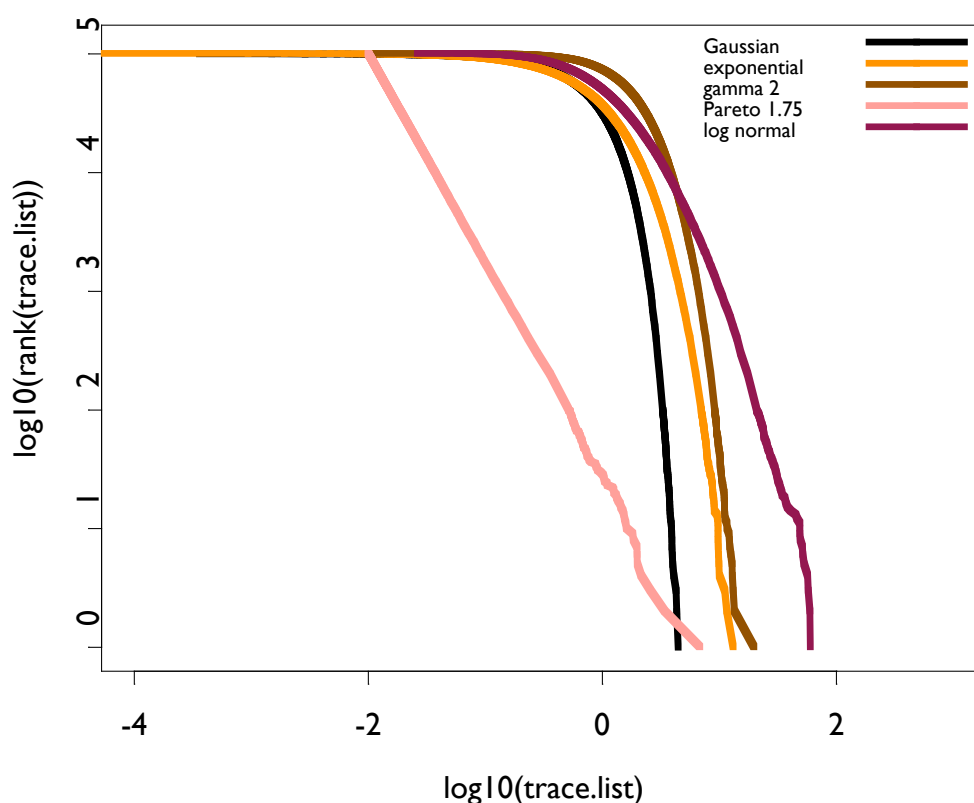


Figure 5.8: Zipf plot of samples from exponential samples compared to the Pareto distribution

The difference between the tails of the various distributions of the exponential family and the straight line of the Pareto is obvious. The more subtle but nonetheless definite differences within the exponential family highlight the attraction of the L -moment system to hydrologists, with their need to accurately estimate the likely occurrence of extreme events.

5.3.1 Power-law tails and rates of profit

Given our investigation of the sample L -moments of Gillman 4, and its importance in testing Farjoun and Machover, we begin with that measure. As in Chapter Four, the data we use in discussing the company-size weighted measure is the concatenation of our 100 RS2 samples.

Figure 5.9 compares size-weighted and unweighted versions of Gillman 4; for purposes of comparison we include the plot of the gamma distribution with shape = 2, previously shown in Figure 5.8. Figure 5.10 presents part of the Zipf plot for each of the five annual estimates of the size-weighted distribution.

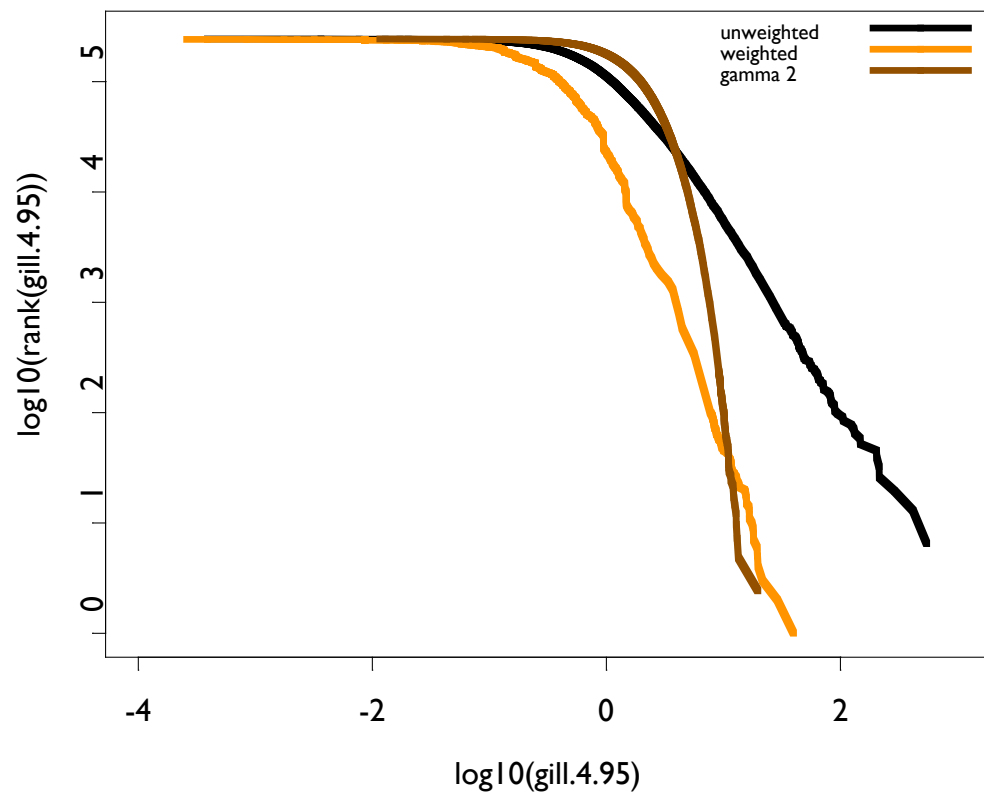


Figure 5.9: Zipf plot of unweighted and weighted Gillman 4 measures, with gamma distribution

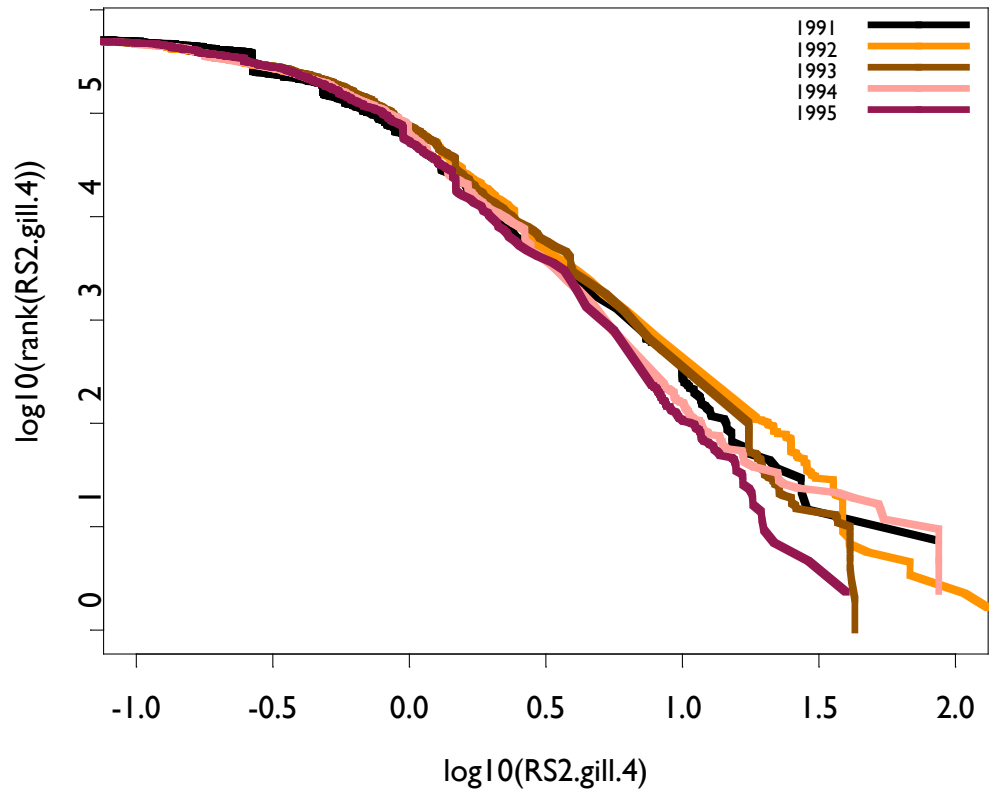


Figure 5.10: Zipf plot of annual variation in weighted Gillman 4 measures

There is a good reason for the similarity of the log normal distribution and the power law tails of the empirical distributions, as Mitzenmacher (2003: 229) points out. Consider the logarithm of the density function of the log normal distribution

$$\begin{aligned} \ln f(x) &= \ln x - \frac{(\ln x - \mu)^2}{2\sigma^2} \\ &= \frac{(\ln x)^2}{2\sigma^2} - \frac{\mu \ln x}{\sigma^2} + \frac{\mu^2}{2\sigma^2} \end{aligned}$$

where μ and σ are the mean and standard deviation of the corresponding normal distribution. With σ sufficiently large, on a log-log plot the logarithm of the density function will appear approximately linear over several orders of magnitude. For comparison, consider perhaps the best-known power law distribution, the Pareto: with shape parameter α and scale parameter k , the logarithm of the density function is exactly linear:

$$\ln f(x) = (\alpha - 1) \ln x + \ln k + \ln$$

As we have seen in Figure 5.8 the Zipf plot of the Pareto is likewise exactly linear (discounting deviations arising from the fact that we plot a sample, not the theoretical curve), and a similar relationship exists between the Pareto and log normal distributions in this case also (although for the log normal the density function is easier to work with).

Both weighted and unweighted versions of the profit rate measure display power law tails. To our knowledge this is the first time this phenomenon has been demonstrated for profit-rate distributions.

The slow rate of change of curvature of the plots makes it difficult to assess the precise start of the power law sections, but it appears that in both cases around 10 per cent of the data has the characteristic power law tail. In the weighted case the approximate range of profit rates included in this tail is from 1 (that is, 100 per cent) to 10 (1,000 per cent), where the estimated mean is 0.45 (45 per cent).

This combination of exponential law for the main mass of data with power law tail(s) is widely accepted as a stylised fact describing not only the firm-size distribution but also the distributions of wealth and income and the returns in financial markets.⁶⁴ On the firm-size distribution, see Sutton (1997) for a survey; more recent contributions include Fujiwara (2004) and Russo, Delli Gatti *et al.* (2006); on wealth and income, see Fujiwara (2003), and a very interesting recent paper by Braun (2006); on market returns see Mantegna and Stanley (2000).

Power law tails are a feature of all our profit rate distributions, as the figures below demonstrate. However, before these inter-profit rate measure comparisons we draw

⁶⁴ Interestingly Stanley *et al.* (1995), who as we have seen above claim to have introduced Zipf plot analysis into economics in connection with this topic, use it to show that with their data the firm-size distribution *does not* exhibit power law characteristics, since the tails are insufficiently heavy.

attention to the right-hand panel of Figure 5.9, comparing the distribution of the size-weighted version of Gillman 4 in different years.

The upper portion of each plot shows that the exponential portion of the data varies from year to year, presumably reflecting annual variation in the shape and scale of the distribution. But it also appears that the slope of (and hence the range covered by) the power law tail undergoes annual change.

Clearly the annual variation in the minimum of the power law tail is comparatively small. Compared to the variation in the maximum value attained it may be taken as zero at a first approximation. That being so, the range covered by the tail varies by about half an order of magnitude – say between $10^{1.6}$ and $10^{2.1}$, or 40 to 125 (profit rates of 4,000 per cent to 12,500 per cent). This supports our conjecture above (section 5.3.2) that it is variation in the tail of the distribution that drives the annual variation in L -kurtosis in the Gillman 4 definition of the profit rate.

We now compare the tail behaviour of different profit rate definitions, in both weighted and unweighted forms, and then examine the annual variations in tail behaviour of selected measures.

5.3.2 Comparison of power-law tails between profit rate measures

We begin by comparing the unweighted profit rate measure definitions, grouped as before into Gillman's marxian and 'capitalist' measures, the accounting ratios, and Glick's measures (in this case we present the Glick measures in two panels to aid clarity in plotting).

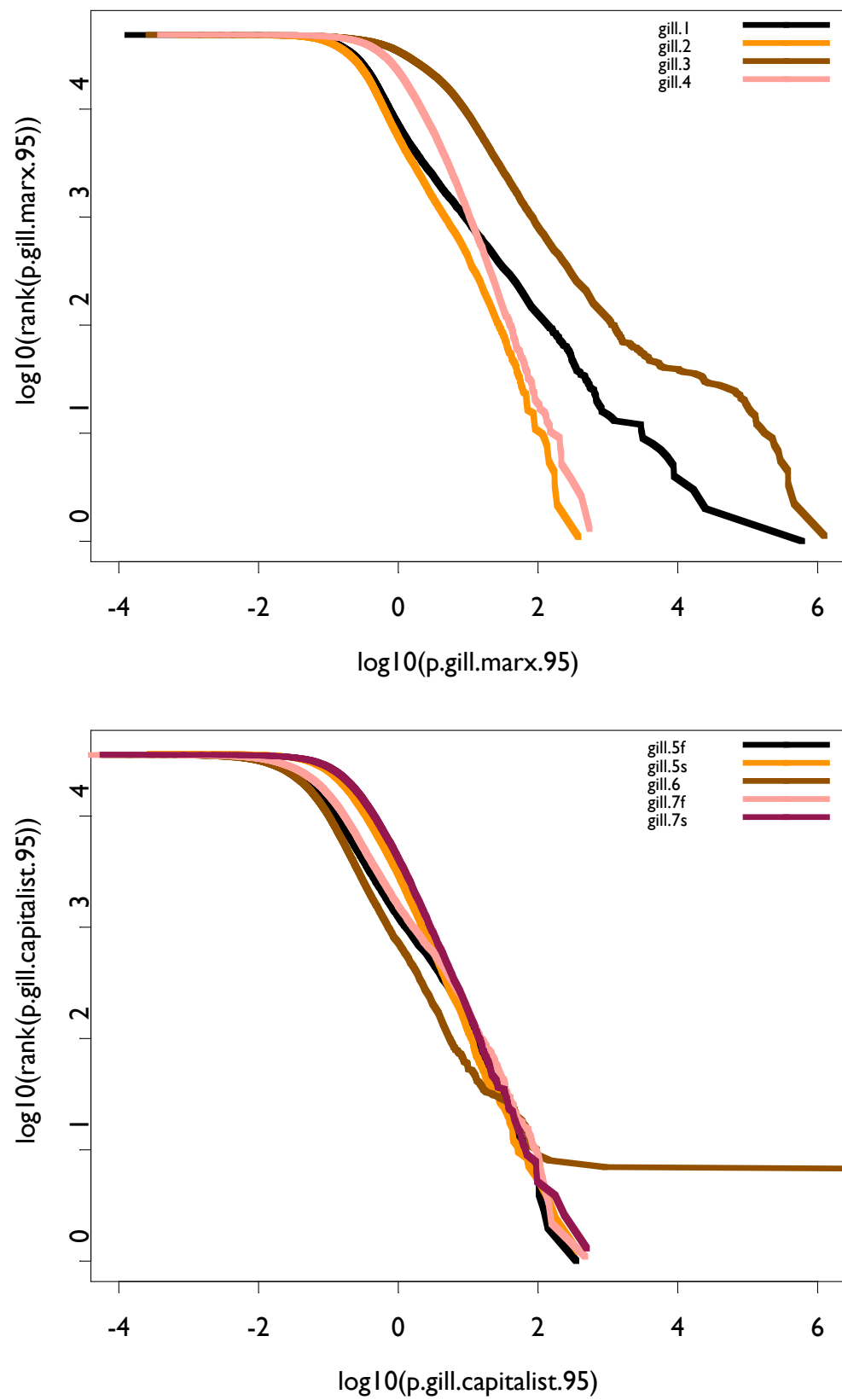


Figure 5.11(a): Zipf plot analysis, unweighted Gillman profit rate measures, marxian and capitalist

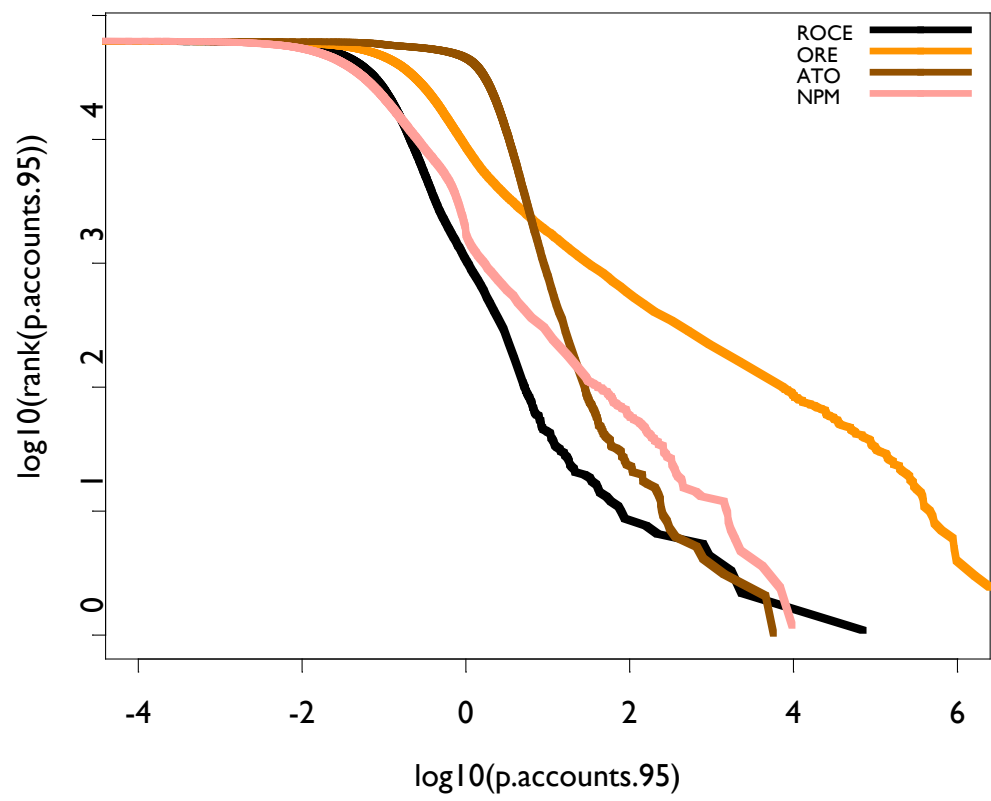


Figure 5.11(b): Zipf plot analysis, unweighted accounting ratios

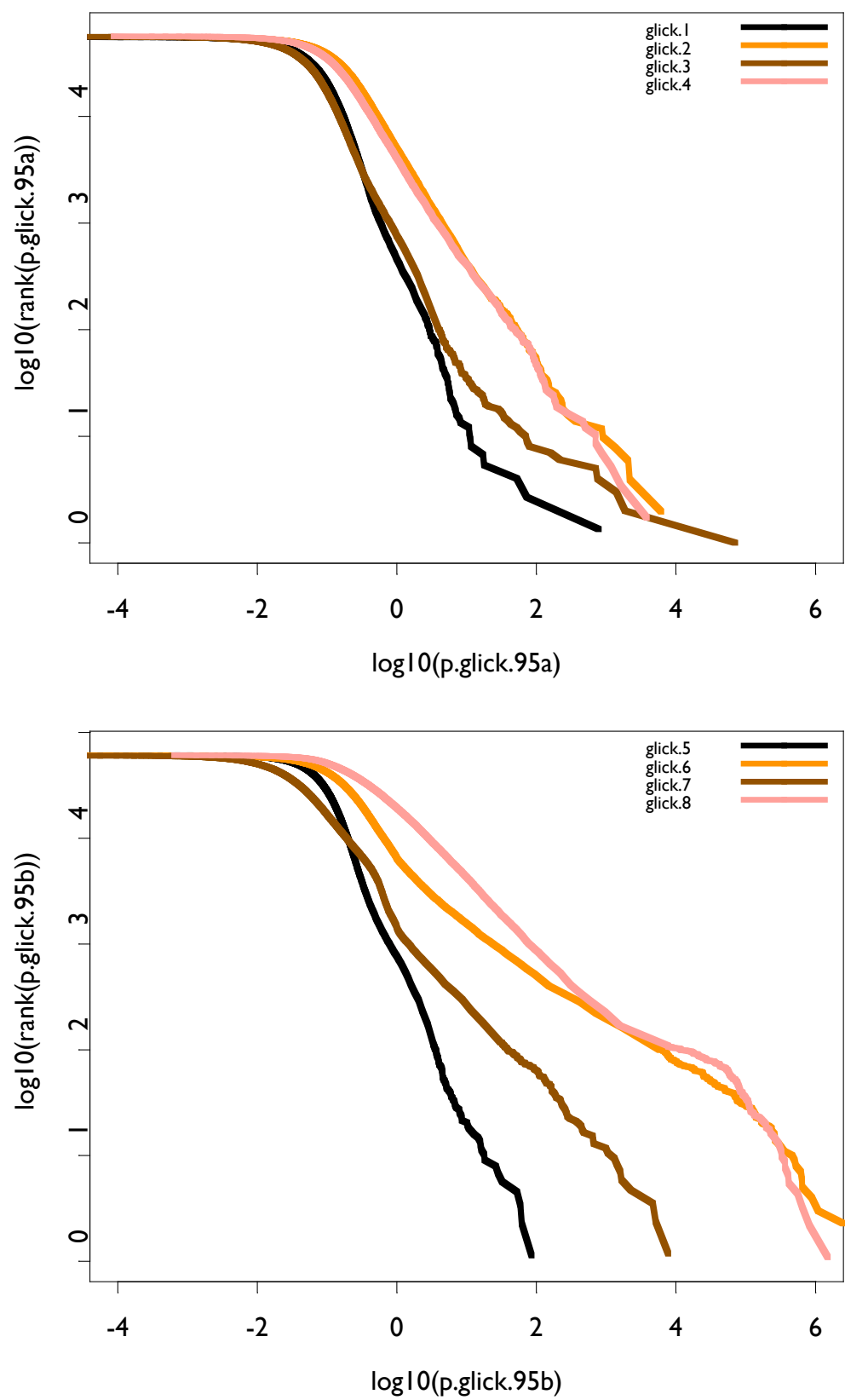


Figure5.1 I (a): Zipf plot analysis, unweighted Glick profit rate measures

In each plot in Figure 5.11 about $10^{1.6}$ observations, or around 40 per cent of the data, are in the power-law tail, although this proportion appears to be somewhat less for the Glick measures (about $10^{1.5}$ observations, or around 30 per cent of the data) – and much less for Gillman 2 and 4: around 10 per cent, as already seen. These results explain the very large skewness and kurtosis – in both conventional and *L*-moment forms – which we saw in Chapters Three and Four.

However, even more striking here are the ranges covered by the tails of certain measures – six orders of magnitude in the cases of Gillman 1 and 3, ORE and Glick 7 and 8, and four orders in the cases of the remaining three accounting ratios and Glick 2 to 4 and Glick 6. These may be contrasted with the two orders of magnitude covered by the power law tails of Gillman 2 and 4, Gillman's capitalist measures (including Gillman 6, bar a handful of extreme observations) and Glick 1 and 5.

A number of intriguing patterns are evident in Figure 5.11, but to avoid the scope of the discussion becoming unmanageable we resist the temptation to comment on all but one. This is the fact that the plots of several measures show first a transition from exponential to power law form, and then a second transition to a shallower slope towards the extreme end of the tail. This is discernable among the accounting ratios and Glick 1 and 3.

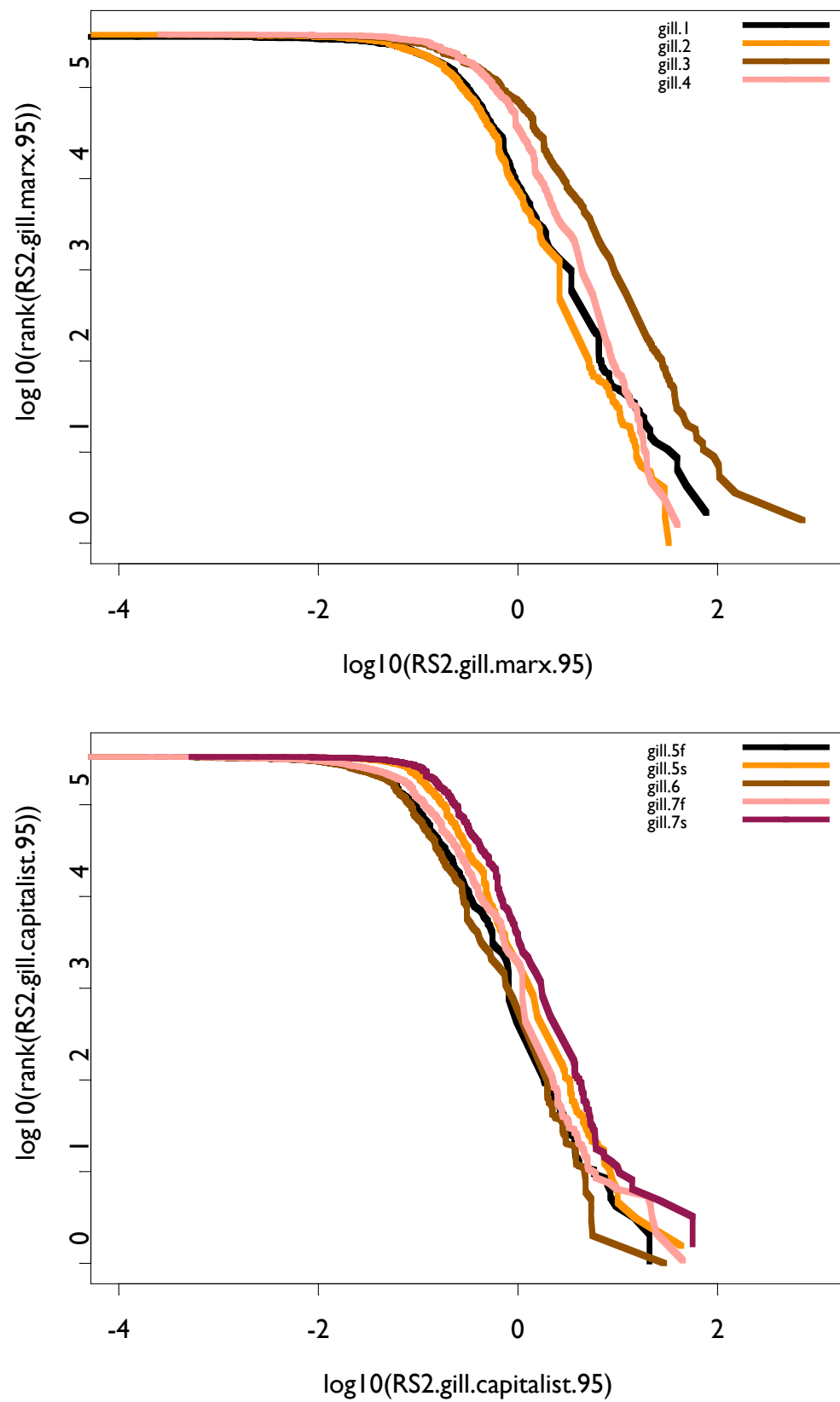


Figure5.12(a): Zipf plot analysis, size-weighted Gillman profit rate measures, marxian and capitalist

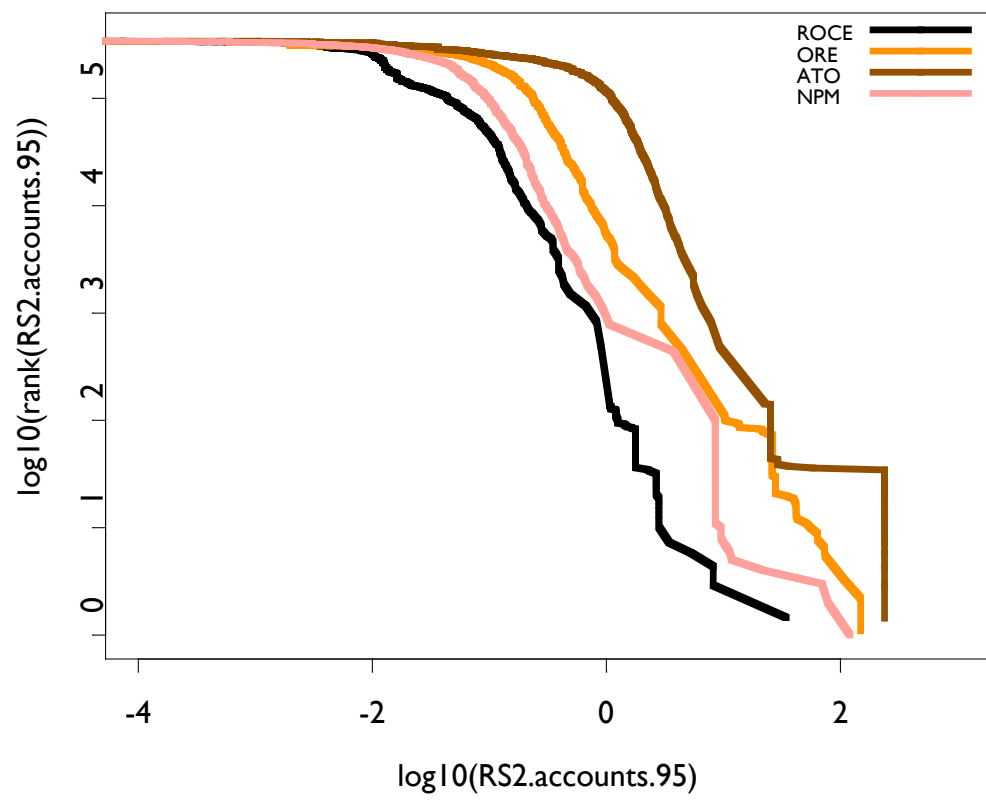


Figure5.12(b): Zipf plot analysis, size-weighted accounting ratios

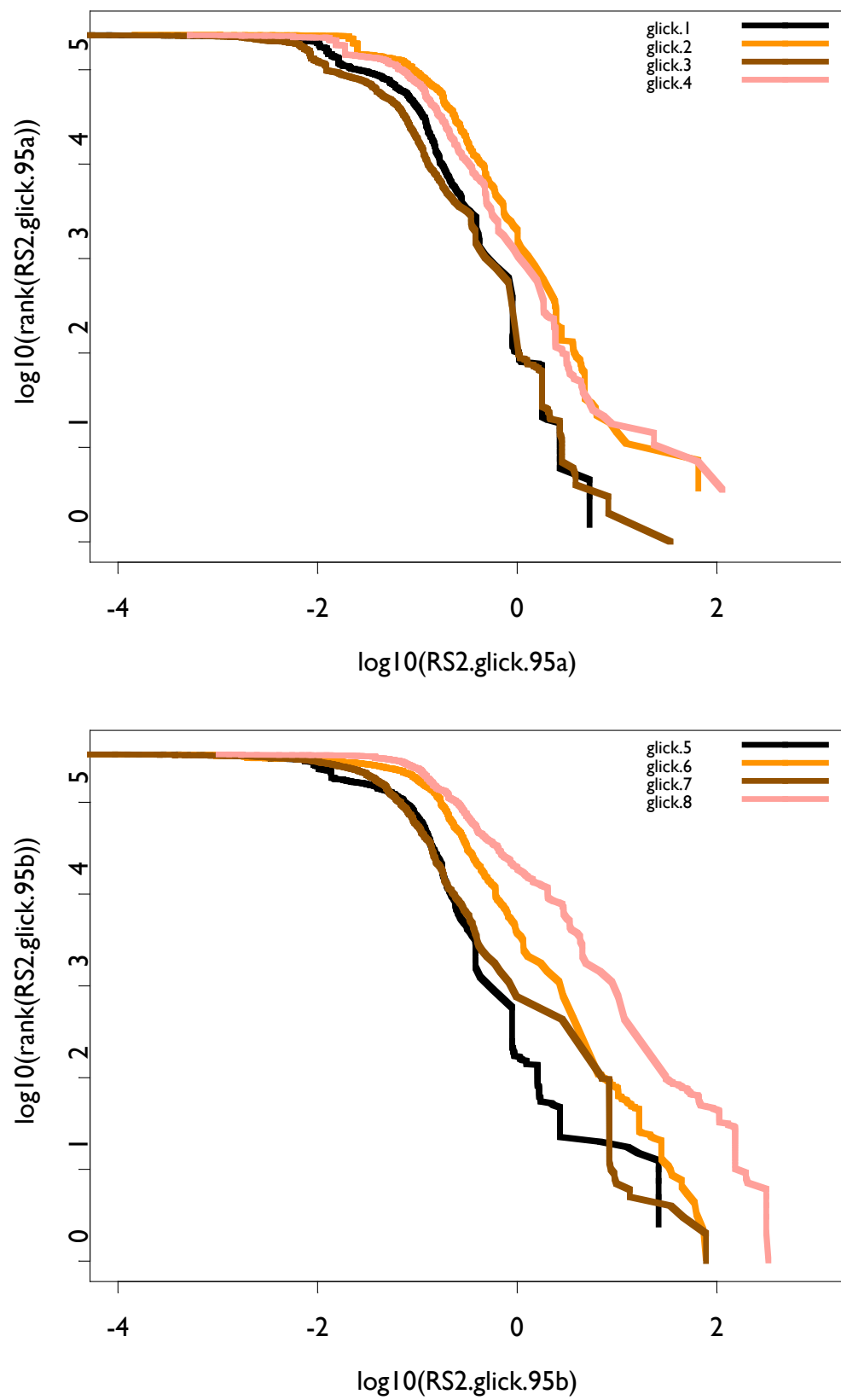


Figure5.12(c): Zipf plot analysis, size-weighted Glick profit rate measures

In Figure 5.12 we compare the weighted versions of our profit rate measures. Three points should be noticed here. Firstly, the RS2 weighting procedure reduces the extent of the tails of all measures and homogenises them, in terms of the number of orders of magnitude covered, as indicated by the steeper and now similar gradients of the power-law portions. (However, the differing lateral displacements of the various tails means that the absolute value of their ranges varies). Second, comparison with Figure 5.11 shows that the exponential part of the distribution has been changed, further undermining our hypothesis that the RS2 procedure might act as a fuzzy rejection scheme for eliminating contaminant observations of non-capitalist entities from the firm-level distribution. Third, the weighting process does not seem to reduce the proportion of the data which they cover: from around 10–17 per cent for the Marxian measures and the accounting ratios, to 40 per cent for Gillman’s capitalist measures and for Glick’s. The implication is that while there is a clear inverse association between size of firm and range of attainable profit rates, it is nonetheless the case that the proportion of *firms* associated with power-law tails is about the same as the proportion of *capital* which is so linked.

Finally, note that comparison of Figure 5.12 with Figure 5.11 reinforces the point that the weighting procedure has not had the effect of a fuzzy rejection scheme for contaminant observations. Hence we cannot now maintain that estimation of capital-level distributions is equivalent to estimation of company-level ones.

5.3.3 Annual variation in power-law tails

In discussing the tails of the size-weighted Gillman 4 measure, illustrated in Figure 5.10, we suggested that the annual variation in the power law tails, expressed in the varying slopes, explained the high variability in the estimated *L*-kurtosis of that measure’s distribution. We now show that similar patterns of annual variation exist in respect of the size-weighted versions of other distributions. We confine our investigation of annual variations in the power-law tails to Gillman’s Marxian measures, the four accounting ratios, and the four measures which Glick found to give the best evidence for his notion of gravitation of profit rates (Glick 1, 3, 5 and 6).

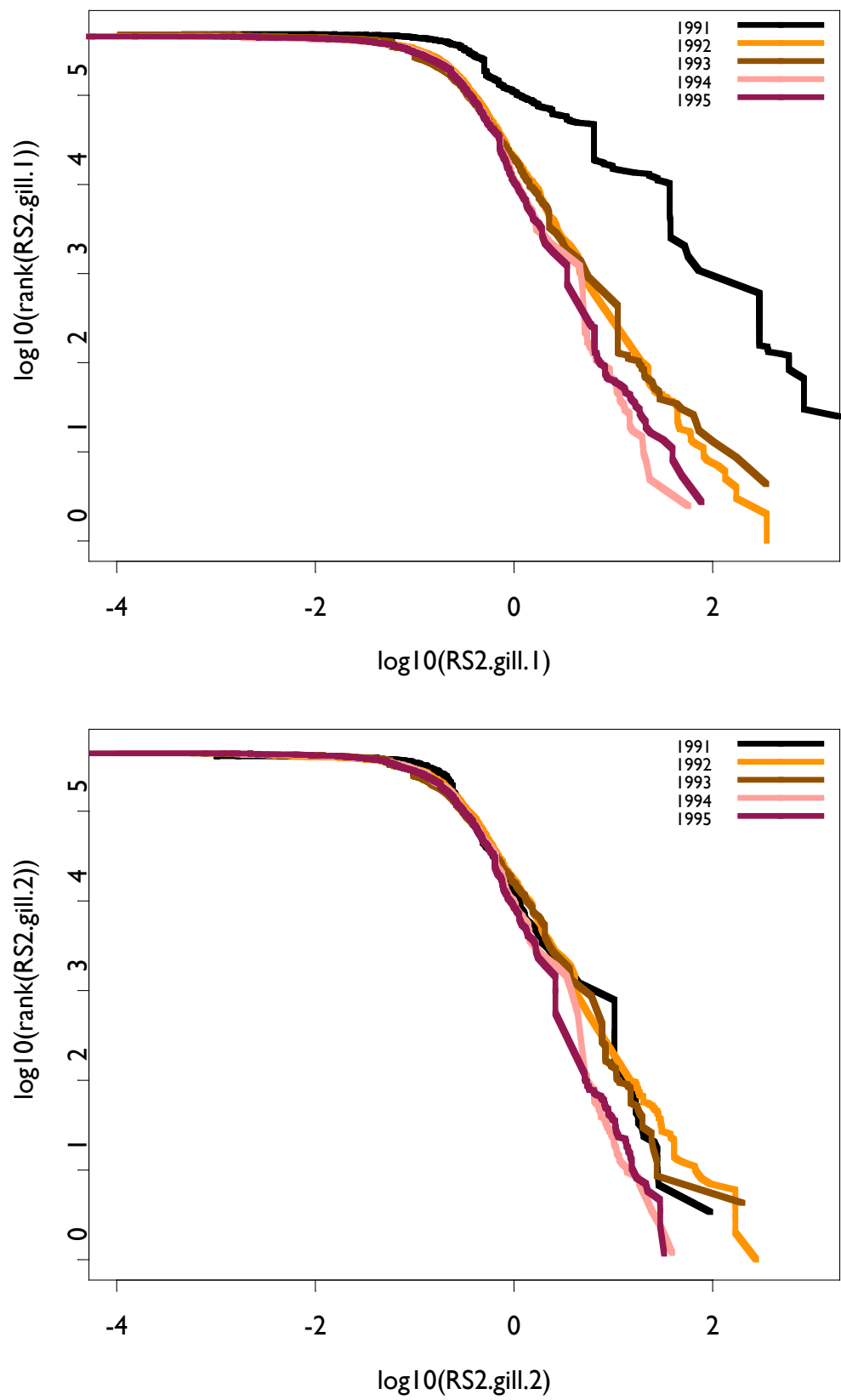


Figure5.13(a): Zipf plot analysis of annual variation; Gillman marxian measures

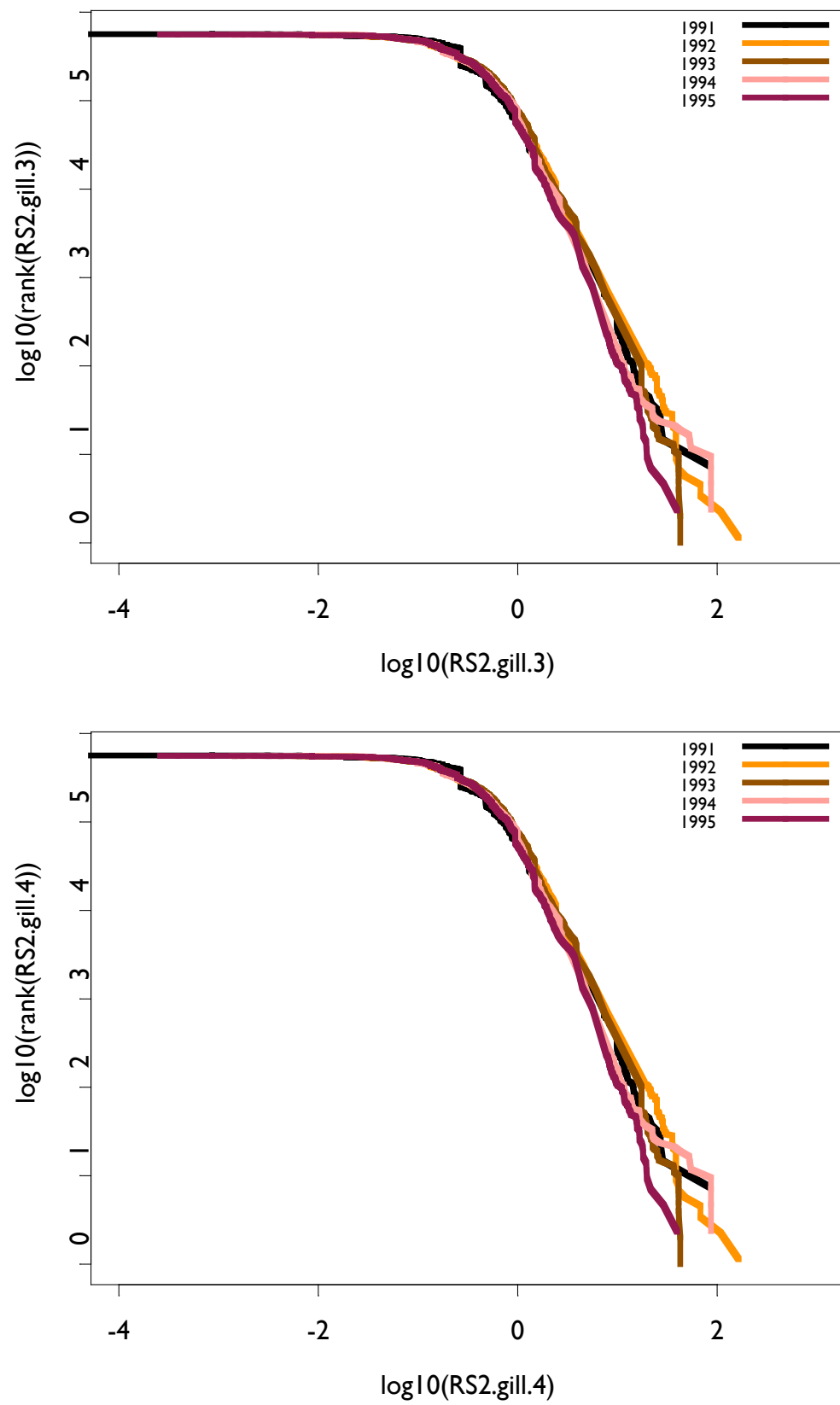


Figure5.13(b): Zipf plot analysis of annual variation; Gillman marxian measures

We begin with Gillman's four Marxian measures, illustrated in Figure 5.13. Very broadly, all these measures have tails whose total range is around two orders of magnitude: a range of profit rates, in percentage terms, from 100 per cent to 10,000 per cent. Variation in the exponential part of each measure's distribution is obscured by the log scaling, but we know from the right-hand panel of Figure 5.9 (which focused on the transition between the exponential and power law parts of the Gillman 4 distribution) that this variation is small in comparison with that of the power law section.

Thus the principal feature of all four measures is the annual variation in tail slope – hence in the value of the power-law exponent (sparsity of available data accounts for the anomalous appearance of the distribution of Gillman 1 for 1991). Thus the maxima of the tails have ranges of the order of about $10^{1.6} = 40$ to $10^{2.2} = 160$ for Gillman 4 (in percentage terms, profit rates from 4,000 to 16,000 per cent).

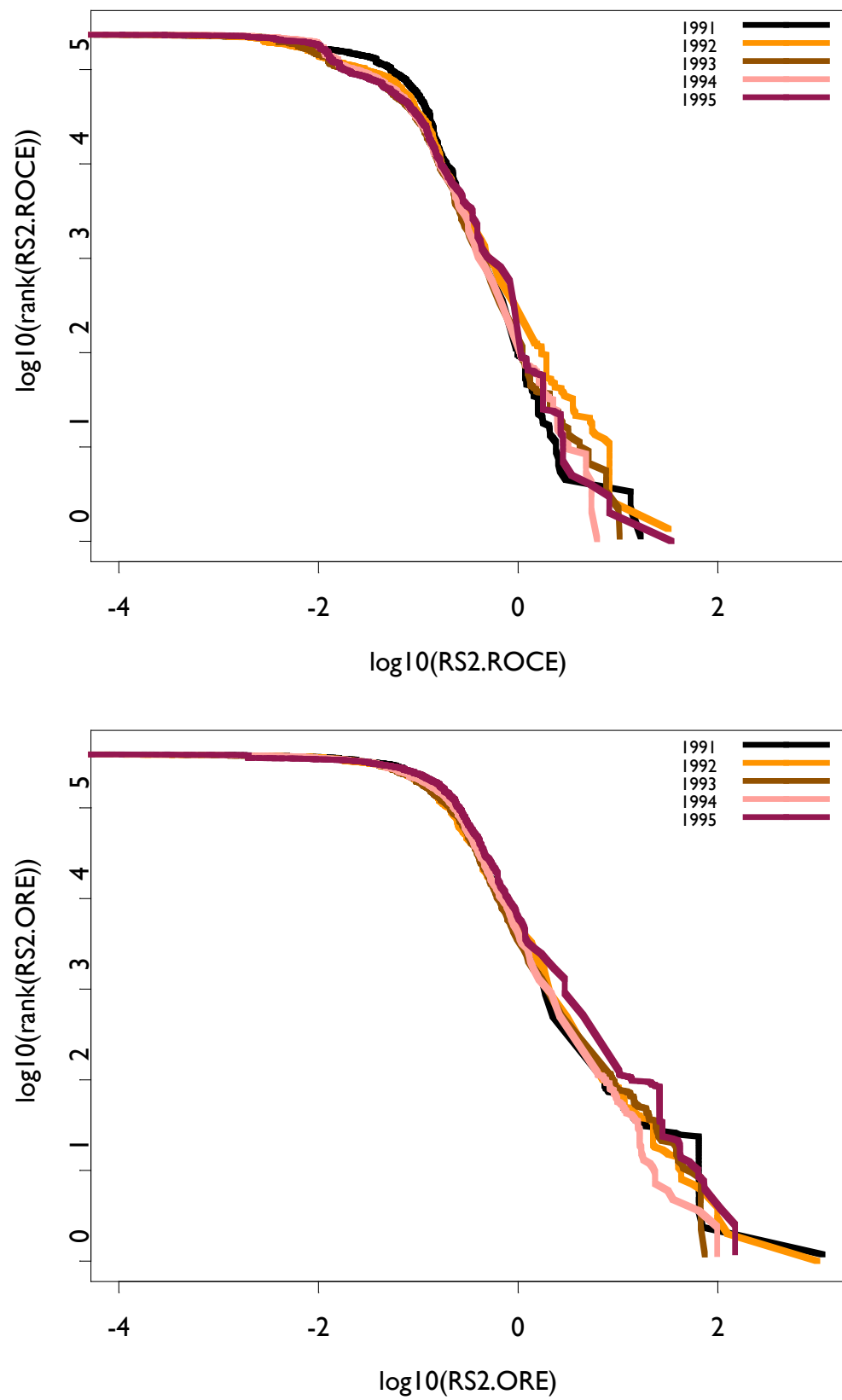


Figure5.14(a): Zipf plot analysis of annual variation; accounting ratios

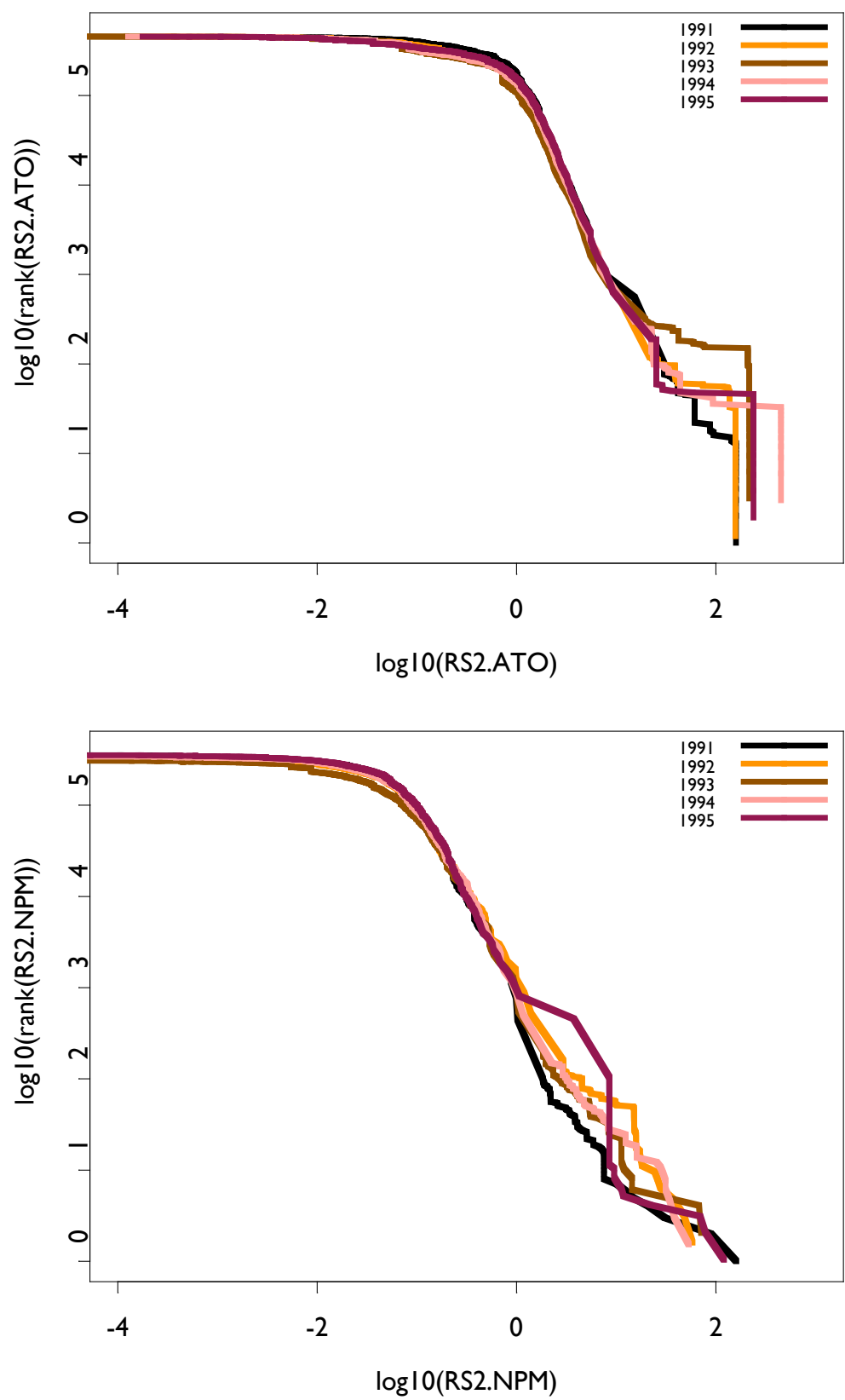


Figure5.14(b): Zipf plot analysis of annual variation; accounting ratios

In Figure 5.14 we show annual variations in the accounting ratios. As with the Gillman measures just examined, the range covered by the tails of these ratios is approximately two orders of magnitude. However, the proportion of the data included in the tails is greater, as already noted.

A distinctive feature here is the way in which the extreme tails of three of the four ratios exhibit either a change of gradient or a marked increase in variability, or both. These are the operating return on equity (ORE), the asset turnover ratio (ATO), and the net profit margin (NPM). (The exception here, ROCE, is also exceptional in showing much greater evidence of annual variation in the exponential part of the plot.)

Not too much importance should be placed on the apparent increase in annual variability in this part of the tails. To see this, note that our weighting procedure effectively produces an estimate of the distribution of the rate of return across units of capital (albeit one with spikes and clumps in the density, due to the clustering of capital in firms, some very large in relation to the total capital). The parts of the tails we are discussing are those between values of 10^0 and 10^3 on the vertical axis – in other words the 1,000 largest values. While these are of course rates of return achieved by companies (since that is the origin of the data), in principle they here represent the rates of return accruing to individual £1 units of capital, hence to the highest-earning £1,000 of the total invested in the corporate sector.⁶⁵ In practice; the approximate size of our concatenated RS2 samples lies between 70,000 data points (Glick 2) and 410,000 (Gill 6).

⁶⁵ About £929bn, in the case of the capital measure used to calculate Gillman 4.

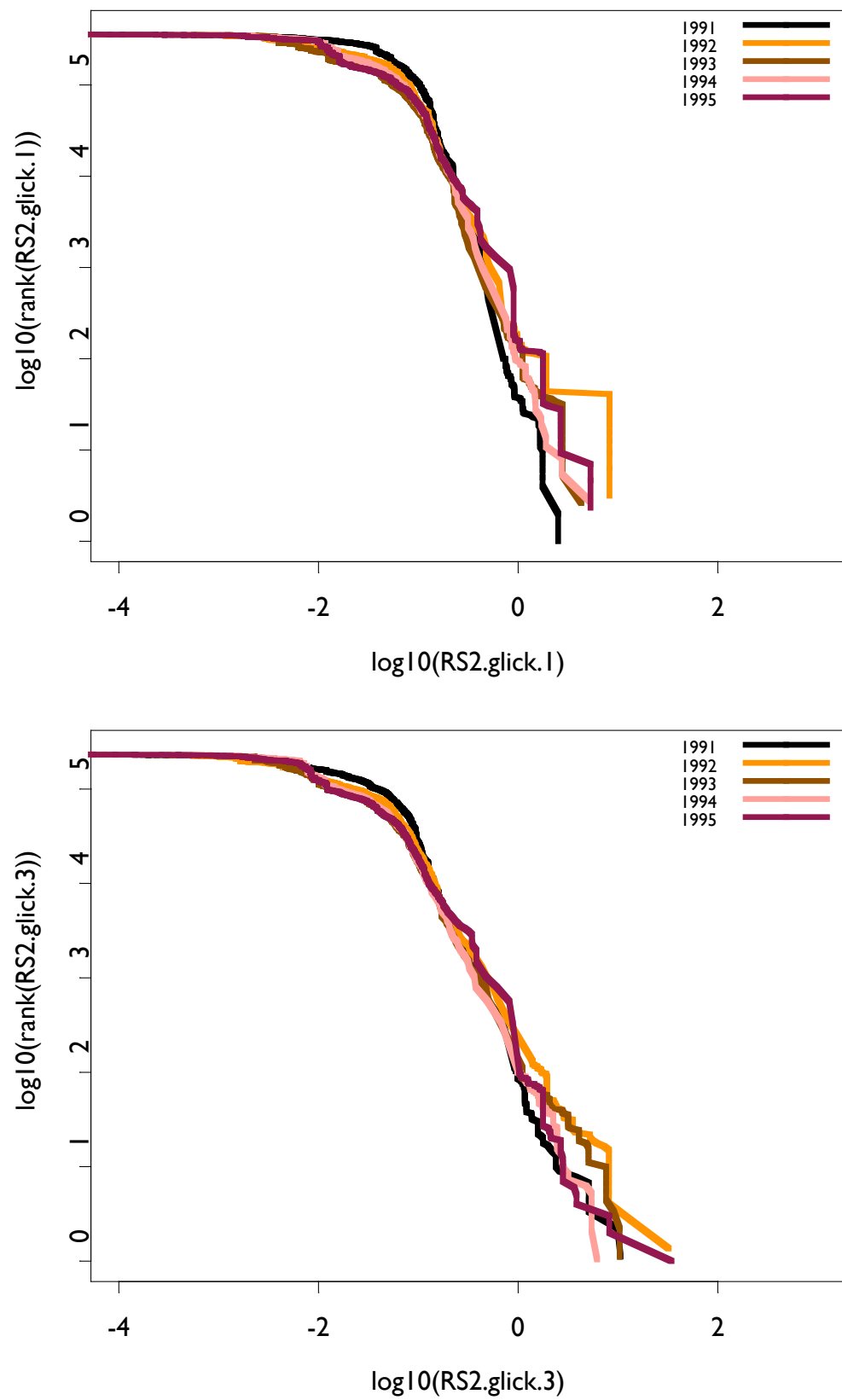


Figure5.15(a): Zipf plot analysis of annual variation; Glick measures

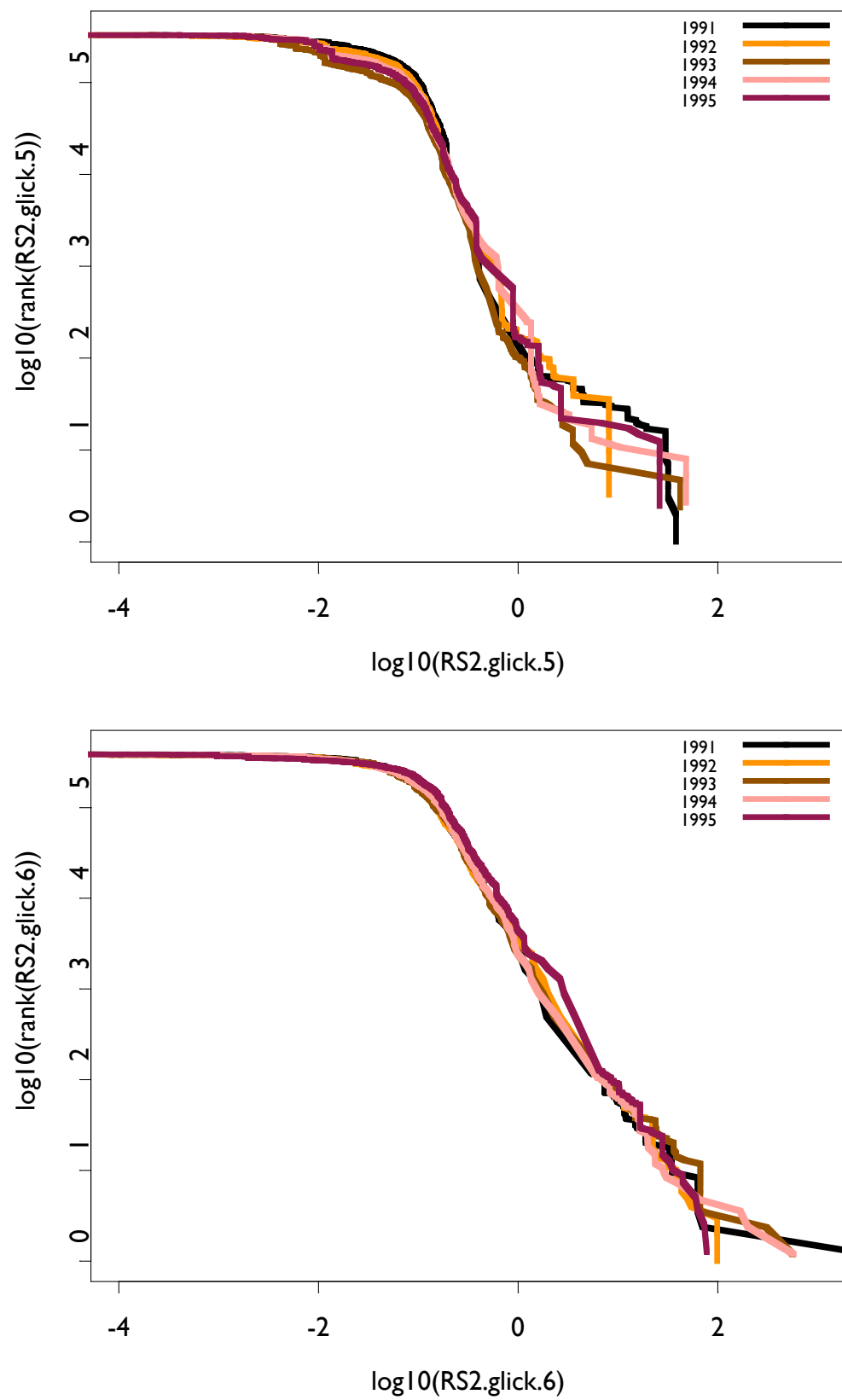


Figure5.15(b): Zipf plot analysis of annual variation; Glick measures

Our final set of Zipf plots, Figure 5.15, shows the tails of our selected Glick profit rate measures. First, we note that the extreme tails of two measures, Glick 5, and to a lesser extent Glick 6, display the same tendency to change of gradient or increased variability, or both, which we saw in the case of the accounting ratios. These three Glick measures (together with ROCE) also share another feature, subjectively greater variation in their exponential sections, compared to Gillman's Marxian measures. Finally, Glick 5 and 6 have tails with ranges somewhat greater than do other measures: perhaps as much as three orders of magnitude in the case of Glick 6 (but note that if one neglects the 100 largest observations in Glick 5, this measure actually has a shorter tail than most other measures, extending only over about one order of magnitude).

One feature of the plots of point clouds in L -skewness, L -kurtosis space which remains to be explained is the differences in the pattern of annual variation in skewness and kurtosis between Gillman's 'marxian' measures (together with ATO) and the remaining measures (excepting Glick 8). Broadly speaking the former group all fell within the band of Hosking and Wallis loci whereas the others lay well above it; with the exception of Gillman 1 the first group displayed less annual variation than the second, perhaps particularly so in respect of skewness.

In the case of Gillman 4 we suggested that the variation in the power law tails explained annual variation in kurtosis; in the case of the non-'marxian' measures we have instead to explain high – but only lightly-varying – kurtosis, combined with wide variations in skewness.

Our suggestion here is that the latter can be explained by variation in the balance between upper and lower tails, unremarked because our Zipf plot analysis has only looked at the former. Consider Table 5.2: this excerpts from Table 3.5 the information about the lower tail quantiles of the *unweighted* profit rate measures.

As can be seen, only the 'marxian' measures (and ATO) have five per cent quantiles which are non-negative (while Gillman 1 has a zero two per cent quantile and Gillman 4

almost does so). Our suggestion here is that both the large variations in skewness and the greater kurtosis are explained by non-symmetric annual variation in the tails of those profit rate measures where the lower tails contain a significant weight of negative values.

Table 5.2: tail quantiles (excerpted from Table 3.5)

PRM	0.0%	0.5%	1.0%	2.0%	5.0%
p.gill.1	-0.996	-0.176	-0.0626	0	0.0392
p.gill.2	-0.996	-0.281	-0.145	-0.0383	0.0312
p.gill.3	-2.23e5	-1.66	-0.372	-0.0573	0.0769
p.gill.4	-1.33e2	-0.346	-0.121	-4.14e-3	0.0851
p.gill.5s	-2.32e2	-3.78	-2.1	-1.11	-0.406
p.gill.5f	-2.1e2	-1.97	-0.96	-0.473	-0.167
p.gill.6	-7.33e2	-0.664	-0.454	-0.276	-0.113
p.gill.7s	-1.86e2	-3.02	-1.7	-0.861	-0.289
p.gill.7f	-1.58e2	-1.31	-0.702	-0.355	-0.117
p.ORE	-1.47e5	-9.62	-4.26	-1.94	-0.684
p.ROCE	-8.41e1	-0.799	-0.509	-0.305	-0.127
p.ATO	-0.0255	8.96e-3	0.0229	0.0445	0.0884
p.NPM	-4.16e3	-2.04	-0.828	-0.384	-0.119
p.glick.1	-8.41e1	-0.803	-0.52	-0.32	-0.137
p.glick.2	-9.01e21	-5.71	-2.88	-1.41	-0.496
p.glick.3	-8.46e1	-0.768	-0.474	-0.278	-0.115
p.glick.4	-9.01e21	-4.73	-2.46	-1.22	-0.419
p.glick.5	-8.44e1	-0.697	-0.424	-0.236	-0.0827
p.glick.6	-1.6e5	-8.7	-3.8	-1.76	-0.634
p.glick.7	-4.16e3	-1.98	-0.816	-0.364	-0.11
p.glick.8	-4.56e5	-2.42e1	-9.75	-3.81	-0.811

Note: shaded cells indicate non-negative quantiles

5.3.4 Summary of Zipf plot analysis

Although our analysis is broadly qualitative rather than quantitative, several conclusions appear to be well-supported. First, the high levels of L -skewness and L -kurtosis found in both weighted and unweighted versions of all the profit-rate measures examined are accounted for by power-law tails, a sign of which is the much steeper slopes of the weighted versions accompanied by reduction in the estimated L -moment ratios.

Second, the variations in L -skewness and L -kurtosis are explained by variation in the length and weight of these tails.

Third, the relatively low weight of the power law tails of the four Gillman Marxian measures accounts for the fact that our RS2 procedure for estimating L -moment ratios is more nearly successful here than in other cases. This is because these tails are not obviously shorter, in the sense of the range of values covered, than those of the other types of measure (although it should be noted that Glick 5, which on one view has shorter tails than the Marxian measures, is the only other measure apart from ATO for which any years' estimates fall within the band of Hosking and Wallis loci).

In the case of Gillman 4, in particular, the upper tail is light enough that it does not significantly interfere with the L -moment ratios' ability to identify a possible log normal distribution.

5.4 Conclusion

In this chapter we began by showing that Farjoun and Machover's probabilistic political economy is entirely in the spirit of Marx's own approach – indeed, we think it little exaggeration to say that their work can best be viewed as a proposal for filling in technical details which the statistical science of Marx's day would have struggled to provide.⁶⁶

We then applied the RS2-L methods developed in Chapter Four in an attempt to test the hypotheses about profit rate distributions put forward by Gibrat and Farjoun and Machover. The results are best described as an encouraging near-miss; Gillman 4, the measure we regard as instantiating the profit rate definition to which Farjoun and Machover's hypothesis relates, was tentatively identified as log normal. As far as it goes, this

⁶⁶ In particular, Maxwell's kinetic theory of gases, the first contribution to statistical mechanics, is an almost exact contemporary of *Capital*: Maxwell first publication in the area was in 1860, and he produced a fully-worked-out statement in 1865, two years before the publication of Volume I.

is direct support for Gibrat, and also weak confirmation of Farjoun and Machover, in view of the fact that the lognormal is a special case of the four-parameter gamma distribution.

An investigation of the sampling properties of the RS2-L procedure suggested that the unsatisfying aspects of the results for Gillman 4 were likely to be the outcome of extensive tails still remaining after the size-weighted sampling process. Removal of discordant samples could be expected to produce more convincing estimates of L -moment ratios.

This was reinforced by the Zipf plot analysis in section 5.4, which showed that both weighted and unweighted versions of this measure have power law tails appended to distributions whose main mass appears to be exponential.

Indeed, all the profit rate measures tested have such power law tails, a fact not previously demonstrated in the literature.

The partial success in the case of Gillman 4 is due to its power law tail being relatively light. Since this profit rate measure is the one we have identified as instantiating Farjoun and Machover's definition of the profit rate relevant to their hypothesis, we have not succeeded in rejecting the hypothesis that other definitions of the rate of profit might also lead to gamma distributions.

Chapter 6 Profit rates and the competitive process

In this chapter we address the conception of capitalist competition as the gravitation of profit rates about a norm, as developed by Mark Glick (1985). This has been seminal in reviving both non-neo-classical ideas about the competitive process and their empirical investigation (see, for example, Christodoulopoulos, 1996, Glick, 1985, Glick and Erbar, 1988, 1990, Glick and Campbell, 1995, Lianos and Droucopoulos, 1993, Maldonado-Filho, 1998, Tsaliki, 1998, Wolff and Dollar, 1992). It has also been warmly endorsed as a conceptual and methodological advance:

Up until a few years ago Stigler's and Brozen's work on [competition and concentration] represented the orthodox opinion, but it was based on limited datasets and primitive econometrics. This opened the door to a more systematic attack on the problem by Mark Glick, Hans Ehrbar and others starting from the Classical point of view of equalization of profit rates using more sophisticated econometrics and uniform, comprehensive data sets.

(Foley, 1989)

Glick argues that the process of capitalist competition has two qualitatively opposite effects. On the one hand, the struggle of individual firms for competitive advantage creates persistent inequality of profit rates *within* industries ('hierarchies' of profit rates, in Glick's terminology). On the other hand, capital flows from industries where the average profit rate is low to those where it is higher, creating a *tendency* for industry average profit rates to equalise over time ('gravitation' of profit rates, in Glick's terminology). However, the gravitational process is constantly counteracted by the ebb and flow of supply and demand for particular commodities and the resulting variations in prices in, and profits of, the industries producing them.

Thus the wide differences between the profit rates of different companies in a given year, and the violent swings from one year to another in the profitability of given firms and industries, hide a more fundamental fact: that the amount by which each industry's profit rate deviates from the economy-wide average in each year should, when averaged over time, be small, and thus that the total of such average deviations be small.

Our argument will be that neither the data nor the methods used in Glick's 1985 thesis permit a genuine test of his theoretical position. From the perspective of the present work our most fundamental methodological objection is that he does not consider any feature of profit-rate distributions other than their variance. However, objections grounded in even the most fundamental principle do not have great force in the eyes of those yet to be convinced of that principle's importance – unless they can also be shown that disregard of the principle has substantive consequences. Thus our aim here is two-fold: first, to show that his methods are indeed problematic (an *internal* critique), and then to apply our own and demonstrate that they provide better insight into the gravitation of industry profit rates (an *external* critique).

The present chapter therefore has six sections: in Section 1 we review Glick's thesis (1985), and in Section 2 we present our critique of this. In Section 3 we use our FAME data set to carry out tests comparable to Glick's work on the US manufacturing sector, and extend them to two-digit industries covering the whole UK economy. This addresses the point that the theory of competition to which Glick appeals only has meaning in the context of the whole economy, and that it is therefore not useful to confine investigation to manufacturing industry alone.

In Section 4 we first test Glick's methods at different levels of aggregation and confirm that his measure of profit-rate gravitation produces the expected results when extended to higher and lower levels of aggregation than the two-digit level. We then test a proposed modification of his gravitation measure; the result is that difficulties found in Section 3 are relieved, but at the cost of introducing a perverse correlation between level of aggregation and degree of gravitation. The final part of this section concludes our internal critique by calculating both modified and original gravitation measures for a range of profit rate measures, including those which our discussion in Chapter Two, section 2, concluded were theoretically preferable to those favoured by Glick. Unlike the latter, our preferred measures not only display the expected correlation between aggregation and gravitation, when tested with our modified version of Glick's test, but produce *greater* evidence of gravitation.

In Section 5 we move to an external critique of Glick using the RS2-L techniques of random-sized random sampling and *L*-moment ratio analysis developed in Chapter Four and tested in Chapter Five. First we show why Glick's own gravitation measure produces different results to our modified version of it. We then demonstrate that his implicit assumption that within-industry distributions are similar cannot be taken for granted, and in fact turns out to be particularly unreliable in the case of Glick's preferred profit rate measures. Section 6 concludes the chapter. In what follows, all citations of Glick not otherwise identified are to his thesis (Glick, 1985).

6.1 Glick on gravitation of profit rates

The aim of Glick's thesis is to substantiate a theory of the competitive process which he attributes to classical political economy – that is, to Smith, Ricardo and Marx.

According to this profit rates within an industry (*intra*-industry profit rates) will be widely dispersed. Firms in the same line of business will employ techniques with varying proportions of fixed to variable capital. Although it is variable capital (that is, labour-power) which creates value, competition enforces a uniform price for identical use-values and thus firms employing more capital-intensive techniques will be more profitable. Moreover this hierarchy of profit rates will show no long-run tendency to gravitation, thanks to the continual search by all firms for ever-more-capital-intensive techniques that will allow them to increase their profit-rate at the expense of their rivals (pages 13–14).

But competition also tends to bring about prices of production which 'correspond to an equal rate of profit for the *average conditions of production* in each industry' (page 13, emphasis in original). This occurs because capital will tend to move from low-profit to high-profit sectors, diminishing the supply and increasing the price of the use-values produced in the former, and vice versa in the latter. Clearly this will not be an instantaneous process, and moreover the demand for given use values will also change as both technology and social needs change, introducing 'perturbation' in the structure of industry profit rates and transforming a tendency towards equalisation of rates into one of 'gravitation around equal centers of gravity' (page 21).

Thus we should not expect to see *inter*-industry profit rates equalise in the short run: Glick specifically criticises previous studies for failing to consider sufficiently long runs of data (pages 106–113). Rather, the aggregate rate of profit across all industries (the general rate of profit) will be an attractor for each individual industry's aggregate profit rate — it will be the overall 'centre of gravity' around which industry centres of gravity (profit rates) 'orbit', in Glick's phrase.⁶⁷

This conception is explicitly contrasted to the neo-classical concern with whether industrial concentration allows individual *firms* to achieve persistently above-average profit rates. In the classical scheme, as described by Glick, firms may or may not benefit from persistent advantages but this is irrelevant, as it is simply a particular case of the persistent hierarchy in intra-industry profit rates. What *would* demonstrate impairment of the competitive process would be evidence that individual *industries* enjoyed persistent advantage.

Testing gravitation therefore involves accepting inter-industry differentiation of profit rates in any given year – indeed, in every year. Meanwhile intra-industry profit rates 'will *most likely also be unequal* because of the stratification of cost structures due to different technologies, economies of scale, *etc.*' (page 20, emphasis added).

⁶⁷ We will examine this phrase in our critique in section 2 below.

To measure the degree of gravitation Glick follows Lévy (1984) in defining three measures of dispersion, $V = V_1 + V_2$, where:

V is the total variance of the industry rates of profit around the yearly means over the complete set of years

V_1 is the total dispersion of the industry long-run average deviations, and

V_2 is the sum, over industries and years, of the total variance of each industry around its own centre of gravity.

Formally, these are defined as follows:

$$V = \frac{1}{T} \sum_i \sum_t k_t^i (r_t^i - \bar{r}_t)^2$$

where k_t^i, r_t^i are the capital and rate of profit of industry i in year t , and \bar{r}_t is the average rate of profit across all industries in year t . Thus V is the average annual sum of (capital-weighted) squared deviations.

$$V_1 = \sum_i \bar{k}^i (\bar{d}^i)^2 \text{ with } \bar{k}^i = \frac{1}{T} \sum_t k_t^i \text{ and } \bar{d}^i = \frac{\sum_t k_t^i (\bar{r}_t^i - \bar{r}_t)^2}{\sum_t k_t^i}$$

where \bar{k}^i is the average share of capital for one industry over the complete period, and \bar{d}^i is the weighted average deviation from the mean in each industry. Finally

$$V_2 = \frac{1}{T} \sum_i \sum_t k_t^i (r_t^i - \bar{r}_t - \bar{d}^i)^2$$

In his Appendix A Glick raises, without definitively answering, the question of in what way these statistics should be normalised – by scale or by level. Consider the three samples (5, 6, 7), (25, 30, 35) and (5, 10, 15): does the first pair have the same dispersion, because one is five times the other (implying normalisation by scale), or the second pair, because the absolute deviations are ± 5 (implying normalisation by level)?

Glick regards V_1 as ‘the best criterion for a measure of classical gravitation’ (page 99): in the case of equal centres of gravity, this term would be zero, since industry centres of gravity would be equal to the average.

Thus, a comparison of the case of equal centers of gravity with large fluctuations with the case of unequal centers with small fluctuations may result in equal V_s , but the first case would produce a small V_1 and a large V_2 , while the second case would record a large V_2 , and a small V_1 .

(page 86, *sic*: presumably the last phrase is a slip for ‘large V_1 , and a small V_2 ’).

Testing any proposition about profitability obviously entails specifying which of a wide range of possible profit rate measures should be used. As discussed in Chapter Two (section 2.2.3), Glick points out that ‘[t]he logic of a rate of profit is a logic of private investment’ (page 87) and argues that those profit rate measures are to be preferred which include financial assets in their measure of capital, on the basis that these represent ‘the total cost advanced in order to generate the income of the firm’ (page 88). He specifically deprecates ‘the “marxist” ratio of total surplus value divided by fixed capital’ (page 91).

Glick’s tests use two different data sets. One is a composite series of industry profit rates for 1969-1982, compiled by the Value Line organisation from the accounts of 1,637 U.S. corporations in such a way as to take account of missing values, and of births and deaths of firms.

Eight profit rate measures are calculated from this, normalised by scale; in Appendix A, 40 further profit rates are computed and the effects of normalisation by scale and level are compared.⁶⁸

⁶⁸ The extra profit rate measures are constructed simply by considering every possible combination of eight measures of profit and six measures of capital, without regard to economic or accounting logic.

The other dataset is taken from the U.S. National Income and Product Accounts (NIPA), and four measures are calculated from it. The NIPA data is only available at the two-digit level of aggregation, and thus for comparability the Value Line series is also aggregated to this level, although the original is at the three-digit level.

An important feature is that the analysis is only carried out in respect of 18 manufacturing industries (see Table 6.4 for details). Glick states (page 59) that this is *‘because of the unsatisfactory initial results we obtain for this data base [NIPA] on the basis of total industries*. Clearly some questionable adjustments would have been necessary to include non-manufacturing industries in the study’ (our emphasis). He does not say here (or, as far as we have been able to discover, anywhere else) in what way the initial results are unsatisfactory, or what the ‘questionable adjustments’ would have to be.

The results obtained from the Value Line data are most nearly comparable with what can be done with our own FAME data set. Table 6.1 shows the values Glick finds for each of his statistics for the eight principal profit rate measures.

Table 6.1: Glick’s Value Line statistics

profit rate measure estimated from Value Line data			V	V ₁	V ₂
Glick 1	ROCE	(profit + net interest)/total assets	0.04243	0.02302	0.01940
Glick 2		(profit + net interest)/(net plant + inventories + cash)	0.10231	0.06495	0.03736
Glick 3		profit/total assets	0.08428	0.04359	0.04070
Glick 4		profit/(net plant + inventories + cash)	0.09395	0.05406	0.03989
Glick 5		(profit + depreciation)/total assets	0.03601	0.02313	0.01288
Glick 6	ORE	profit/equity	0.07205	0.03021	0.04184
Glick 7	NPM	profit/sales	0.12573	0.08726	0.03847
Glick 8		(profit + net interest + taxes)/net plant	0.11887	0.08022	0.03866

Recall that Glick regards V₁ as the best measure of gravitation, and Glick 1 as the most appropriate measure of the profit rate. Table 6.2 shows the value of this for each profit rate measure in rank order; as can be seen, Glick 1 indeed ranks first in degree of gravitation: its V₁ score of 0.02302 indicates the absence of persistent long run deviations of industry profit rates from their centre of gravity. In contrast, the ‘marxist’ measure Glick 8, regarded as the

least likely candidate to exhibit gravitation, comes seventh, a little way ahead of Glick 7 (the profit margin).

Table 6.2: rank order of Glick's Value Line results by V_1

profit rate measure estimated from Value Line data			V_1
Glick 1	ROCE	(profit + net interest)/total assets	0.02302
Glick 5		(profit + depreciation)/total assets	0.02313
Glick 6	ORE	profit/equity	0.03021
Glick 3		profit/total assets	0.04359
Glick 4		profit/(net plant + inventories + cash)	0.05406
Glick 2		(profit + net interest)/(net plant + inventories + cash)	0.06495
Glick 8		(profit + net interest + taxes)/net plant	0.08022
Glick 7	NPM	profit/sales	0.08726

Glick thus finds good accord between his theoretical views and the results produced by his methodology: 'the measure chosen has an important influence on the degree of dispersion of industry long-run centers of gravity' (page 126).

However, it must be noted that the V_1 score for Glick 1 is virtually identical to that of Glick 5; this measure is a variant of Glick 3 (conceptually equivalent to the operating return on equity, ORE), described by Glick as 'a very traditional' measure which he includes for comparison with 'more theoretically specified' ones.

Glick 3 itself comes some way behind Glick 6 (conceptually equivalent to the operating return on capital employed, ROCE), the measure most frequently used in the empirical literature reviewed by Glick, and regarded by some as the measure most likely to exhibit strong gravitation.

6.2 Critique of Glick

A number of reservations must be made to Glick's claim to have satisfactorily tested his hypothesis about profit-rate gravitation. In the course of our review of profit rate measures (Chapter Two, section 2) we pointed out that neither his decision to examine manufacturing industry alone nor his preferred measures of the rate of profit appeared to be in conformity with the tradition in which he places himself.

In this section we will address some further issues which appear questionable even within Glick's own perspective. It will thus be an *internal* critique for the most part, although the last two points also prepare the way for our *external* critique.

This external critique is that put forward by Farjoun and Machover (1983) and discussed by us in Chapter One; that to regard competition as a process of equalisation of profit rates is akin to regarding gravity as an equilibrating force when it pushes the pendulum in one direction, and as a disturbing force when it pushes it in the other. We argued in Chapter Five that Farjoun and Machover's perspective is in accord with Marx's probabilistic and statistical outlook.

On the other hand Glick, and many of his readers, regard his conception of industry profits 'orbiting' the centre of a supposed gravitational system as being derived from the full tradition of classical political economy, including Marx. Debating whether Marx's views are a development of, or a break with, the outlook of Smith and Ricardo would take us outside the aims of the current project. However, the word 'orbit' is not found in any relevant context in any of Glick's three classical sources, and the extent to which any of them employed metaphors of gravitational processes in a way consonant with Glick is highly questionable.⁶⁹ For a recent assessment of the metaphor of profits orbiting a centre of gravity, see Freeman (2006).

Our critique here is directed at the use Glick makes of data derived from company accounts; we do not address that part of his work, using national accounts data, aimed at discovering the period over which gravitation might be expected to be observed. The points we make are as follows.

⁶⁹ Ricardo uses 'gravitation' just once, in Chapter Six of his *Principles*: 'The natural tendency of profits then is to fall; ... This tendency, this gravitation as it were of profits, is happily checked at repeated intervals by the improvements in machinery, connected with the production of necessities'. Although the overall train of thought is clearly that followed by Glick, in Ricardo 'gravitation' refers *only* to the tendency of profits to suffer a secular decline, which is 'checked' by innovation. In other words gravitation is a one-way tendency in Ricardo, whereas it is a two-way one in Glick.

6.2.1 Analysis is at too high a level of aggregation

Although this criticism is made by Glick himself (page 60), it is a passing remark in his description of his NIPA data set and its possible importance requires some effort to extract from his discussion.

In describing his methodology (pages 83–105) he does not provide a detailed discussion of what level of aggregation would be desirable, or in what way the level of aggregation used might bias his results. However, in his discussion of previous studies of industrial concentration (pages 29–53) he reviews a study (Gale and Branch, 1982) which used information from 200 large companies reporting on 2,000 separate businesses, and remarks that ‘[d]ata at the business level allows a very precise study of the product-line level, a concept very close to that of theoretical market.’ (page 51, *sic*: ‘the theoretical market’? ‘theoretical markets’?). Moreover, a subsequent publication of Glick’s which compares gravitation in the U.S. and several European countries using two-digit data for 13 industries states that ‘[o]ur study of profitability differentials will be restricted to a set of rather aggregated industries. Indeed, it would be more desirable to use *an economically more adequate definition* of industry. But this *would require greater disaggregation*.’ (Glick and Erbar, 1988: 182, *emphasis added*).

The data used by Gale and Branch was not the Line of Business data set compiled by the Bureau of Economics of the US Federal Trade Commission but the PIMS dataset.⁷⁰ However, it is worth noting the Bureau’s comments (in its Report on the 1974 Line of Business data) on the question of aggregation. On the one hand, it points out, the Bureau of the Census extends the SIC system to five-digit product classes and seven-digit products (Bureau of Economics, 1981: 3, footnote 2), and ‘[m]any economists accept the four-digit SIC or five-digit Census of Manufactures as the general levels that most closely correspond to economically meaningful markets’. On the other hand, the average number of four-digit

⁷⁰ PIMS (Profit Impact of Market Strategy) is a decision-support system originally developed at the General Electric company and later at the Harvard Business School, and now the core product of a network of management consultancies.

SIC industries in which the largest 200 manufacturing firms was engaged in 1950 was 12.6, while 10 of these firms were engaged in more than 30 industries (for 1968 the diversification is even greater: 39 of the top 200 were engaged in more than 30 industries) (page 2).⁷¹

Glick's favourable reference to the product-line data suggests he might find even four-digit industry data insufficiently disaggregated for tests of gravitation. On the other hand, they will presumably be preferable to those at a higher level of aggregation (because they more closely correspond to the markets for individual commodities). Indeed, even company-level data has to be regarded as aggregate data, and moreover at an unknown and varying level of aggregation from firm to firm. The FTC's comments also suggest that data produced by aggregation from company accounts will be an imperfect substitute for that aggregated from line-of-business surveys (because of the preponderance of large companies, and those companies' extensive diversification).

In fact, the difficulty is shown by the following example (to take just one reason for inter-industry profit differentials): at any given point in time the existence of novel products on the one hand and obsolescent ones on the other will mean industries with respectively higher- and lower-than-average rates of profit. Even within a single period, aggregating numerous high- and low-profit product lines into a smaller number of more widely-drawn industries will reduce the measured inter-industry dispersion. As obsolete products disappear entirely, and transfer of capital undermines the higher profitability of novelties, the dispersion of product-line profit rates will fall, and the more so the longer the period examined – but whatever dispersion remains will still be reduced by aggregation.

Since Glick's test of gravitation is the total of the average annual deviation of industries, testing at a higher level of aggregation must tend to increase the perceived degree of

⁷¹ Ravenscraft (1983) is a traditional structure-performance study using FTC Line of Business data.

gravitation.⁷² Hence an apparent demonstration of gravitation at the two-digit level is always open to being undermined by a demonstration that it is weaker at the four-digit level. Thus testing at the lowest feasible and theoretically-relevant level of aggregation is necessary, since testing at any higher level is vulnerable to the complaint that any gravitation demonstrated is a mere artefact of aggregation.⁷³

If the foregoing holds it is hard to resist the conclusion that working at the 2-digit level cannot show anything of any real relevance to the theory, and arguably even the 4-digit level is still some way off a reasonable approximation. At all events it is clearly important to investigate the effects of performing Glick-type analysis at different levels of aggregation.

6.2.2 The preferred statistic for measuring gravitation is questionable

Although Glick asserts (page 86, page 99) that the absolute size of his statistic V_1 is the best measure of long-run gravitation, it can be argued that a better one is the statistic $\hat{V} = V_1/V$, the *proportion* of V which is accounted for by V_1 . This is on two related grounds: (i) his reasoning in favour of V_1 is arguably incomplete and (ii) it does not provide a criterion for judging whether the degree of dispersion exhibited in a single case in fact demonstrates gravitation.

First, recall his justification for V_1 : ‘comparison of the case of equal centers of gravity with large fluctuations with the case of unequal centers with small fluctuations may result in equal V s, but the first case would produce a small V_1 and a large V_2 ‘ and *vice versa*.

⁷² To make the point most forcefully, consider testing at the 0-digit level – in other words, calculate the deviation of the general rate of profit of the whole economy from itself. Since this is zero one must necessarily find perfect ‘gravitation’.

⁷³ For more on the possible difficulties of aggregation, see the recent controversy on this topic between Cockshott and Cottrell, collaborators who share the Farjoun and Machover perspective, and Kliman (Cockshott and Cottrell, 2003, Kliman, 2002, 2003).

Now suppose the contrary situation: two cases with equal V_1 but unequal V and V_2 . If $V_1^x = V_1^y = 0$ then on any interpretation we have equally perfect gravitation in both cases and further comparison is redundant. But what if $V_1^x = V_1^y \neq 0$?

The two best instances of gravitation, according to Glick's results, are just such a case: he finds $V_1^{\text{Glick 1}} = V_1^{\text{Glick 5}} = 0.023$, while $V^{\text{Glick 1}} = 0.042$ and $V^{\text{Glick 5}} = 0.036$ (and thus $V_2^{\text{Glick 1}} = 0.019$ and $V_2^{\text{Glick 5}} = 0.036$). If we take absolute size of V_1 as the criterion then we again have to say that these are cases of equal gravitation. Yet for Glick 1 the average dispersion of industry profit rates accounts for a little over half (54 per cent) of the total variation, yet for Glick 5 it is nearly two-thirds (64 per cent) of the total and – put like this – it seems odd to say that Glick 5 is as good an example of gravitation as Glick 1.

This brings us to the second point: V_1 measures the total average annual deviation of industry profit rates. $V_1 = 0$ implies perfect gravitation; if V_1 accounted for 100 per cent of the total deviations V one would presumably say that there was no evidence of gravitation. The question thus arises, how big can V_1 be in relation to V and still constitute evidence for gravitation? It would be desirable to have a criterion based on argument from first principles, but even in the absence of such an argument it seems incongruous to suggest that average industry deviations which account for more than half the total count as such evidence.

Taking \hat{V} as our measure of gravitation allows us to declare Glick 1 to demonstrate more gravitation than Glick 5, a conclusion which Glick might welcome, given that Glick 1 is the measure he prefers on theoretical grounds, whereas Glick 5 is merely a variant of a measure (Glick 3) which he regards as theoretically inadequate.

However, use of \hat{V} suggests a radically revised overall ranking, as shown in Table 6.3 below:

Table 6.3: Glick Value Line results ranked by relative gravitation

profit rate measure estimated from Value Line data			V_1/V
Glick 6	ORE	profit/equity	0.4193
Glick 3		profit/total assets	0.5172
Glick 1	ROCE	(profit + net interest)/total assets	0.5425
Glick 4		profit/(net plant + inventories + cash)	0.5754
Glick 2		(profit + net interest)/(net plant + inventories + cash)	0.6348
Glick 5		(profit + depreciation)/total assets	0.6423
Glick 8		(profit + net interest + taxes)/net plant	0.6749
Glick 7	NPM	profit/sales	0.6940

Now the measures which do best are those akin to traditional accounting ratios: Glick 6 (equivalent to the operating return on equity, ORE) and Glick 1 (equivalent to the operating return on equity, ROCE). Moreover, use of \hat{V} together with our proposed quantitative criterion suggests that Glick 6 is the only one that provides reasonably convincing evidence for gravitation (although Glick 3 also does fairly well).

The only conclusions which are unaffected concern the relatively poor results from the ‘marxist’ measure Glick 8 and the profit margin, Glick 7; however, it can now be added that these measures, even considered in isolation from others, provide only weak evidence of gravitation.

6.2.3 The full implications of the hypothesis are not tested

Glick’s approach at least implies that intra-industry profit-rate dispersions (in the sense of the variance of firms’ rates of profit about their industry’s average) should be greater than inter-industry dispersions (in the sense of the variance of industry rates about some overall average).

Admittedly, his only specific comment (page 20) is that the intra-industry profit-rates will ‘most likely’ be ‘unequal’, whereas the inter-industry ones are certainly expected to be. But elsewhere he lays great stress on the expectation of ‘hierarchy’, of ‘stratification’, of intra-industry profit rates – and furthermore on its persistence through time, unlike that of inter-industry profit rates – and it seems reasonable to suppose that the tendency should be

for the former to be greater than the latter. However, Glick's data does not allow him to investigate this.

6.2.4 No attention is paid to the *shape* of profit rate distributions

Although Glick pays passing attention to the question of higher moments of profit rate distributions his remarks are not only extremely cursory but opaque (page 71 ff). In a five-page section titled 'Statistical moments in the pattern of industry rates of profit' he presents a chart showing the frequency of industry mean rates of profit over the entire period covered by his Value Line study (Figure 4.4 on page 75, redrawn as Figure 6.1 below).

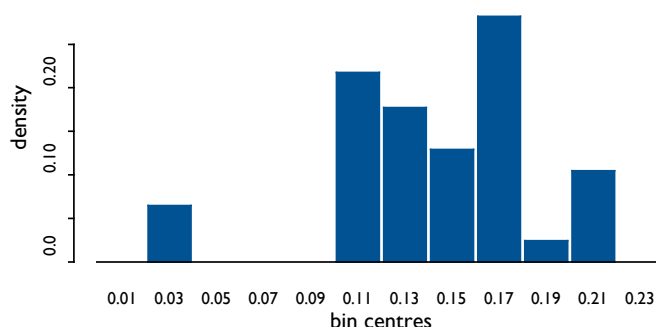


Figure 6.1: histogram of industry average rates of profit, after Figure 4.4 in Glick (1985)

Taking the underlying data, Glick computes the average of the *unweighted* means as 0.147 and the median as 0.146, which implies a skewness of 0.073, computed as $3(\bar{X} - \text{median} / \text{standard deviation})$ (page 74).⁷⁴

However, the skewness weighted by each industry's average share of the total capital stock is 0.85, which Glick says 'is still in a range considered relatively normal'. As reference to Figure 6.1 shows, this distribution is not only not symmetrical, but is not even uni-modal: in fact it has no less than *four* modes and three anti-modes.

⁷⁴ This is the empirical formula due to Doodson (1917).

This thought might inspire two possible strategies. One strategy would follow the line implied by his own discussion, and investigate the higher moments of the distribution of industry average profit rates. The second would note that Glick is wholly dependent on the use of the mean to characterise industry profit rates, and that use of this to test convergence is dependent on the implicit assumption that all industries have intra-industry distributions of a similar kind.

It is easy to see the possible dangers of this by considering the case of two industries with similar means, but where one industry's intra-industry distribution is gamma with shape parameter equal to one (that is, an exponential distribution) and the other industry's is gamma with shape parameter equal to 10 (that is, an approximately normal distribution). The mean and mode will be approximately equal in the second industry, but in the first the modal profit rate will be zero and the median about 0.67.⁷⁵ If these different distributions persisted through time, how realistic would it be to speak of their profit rates equalising? Conversely, but equally problematically, consider the case of two industries with identical modal rates of profit, but with distributions of opposite skewness, and hence different means: does one feel justified in saying that the profit rates of these industries are *not* in some sense convergent, if these distributions persist through time?

6.3 Reproducing Glick

6.3.1 Methodology

We begin our tests by applying Glick's methods to our FAME dataset to provide a baseline for further discussion. We thus calculate his statistic V_1 for the four profit rate measures which he found to exhibit the greatest evidence of gravitation.

As discussed in Chapter Three, limitations on our FAME database mean some profit rate measures can be computed only for the period 1991-1995. These limitations do not apply in the case of the Glick profit rate measures investigated here, but we nonetheless restrict

⁷⁵ Value of exponential median based on Doodson's formula.

our work to these years so as to produce results comparable with those for other profit rate measures which we will examine later in the chapter.

Glick is extremely critical of tests for profit-rate equalisation which do not cover a sufficient time span. Yet while his NIPA data covers 22 years (1958-79), quite long enough to meet his objections, Glick does not hesitate to draw conclusions from his Value Line data, which only cover 14 years (1969-82), a period only slightly longer than the business cycle, which is the source of the problems he identifies in preceding studies. As we shall see in the following section, this short run of data is from some viewpoints a positive advantage.

At this point it is convenient to introduce some extra notation. Glick's decomposition of the total variance V can be used to measure the variance of averages of any entities about some unifying average (V_1), with the variance of the entities about their own averages left as a residual (V_2) making up the overall variation V .

One may usefully distinguish (a) variance of 'industry' (meso-level entity) averages about a common macro average (that is, the economy as a whole), (b) variance of 'firm' (micro-level entity) averages about the averages of 'industries', and (c) variance of firm averages around the macro average. To avoid confusion while at the same time preserving comparability with Glick, we thus propose the notation V^a , V^b and V^c (with appropriate subscripts and hats where necessary) to denote respectively meso/macro, micro/meso and micro/macro situations.

This is a scheme in which Glick's own quantities do not have a clear place (inasmuch as they concern the variance of lower-meso-level objects about higher-meso ones – that is, comparison of 2-digit industry averages to the manufacturing average) although his own presentation implies that his results are type V^a . We therefore consider those tests which are strictly analogous to his alongside V^a tests proper.

Notwithstanding their formal similarity, these three variants of Glick's V -statistics measure different things. Type V^a statistics relate to industries, and thus V_1^a measures

what Glick (1985) wants to measure, namely the persistence of inter-industry profit differentials.

Type V^b statistics relate to firms: thus V_1^b measures the persistence of individual firms' profit differentials within a given industry. However in Glick's theoretical perspective this is *not* a measure of market power due to concentration of industry, since persistent variation in profit rates implies some firms experiencing persistently *lower* profit rates than average (unless one thinks, which the standard literature does not, of the advantages of market power accruing to the beneficiaries at the expense of other firms within the industry concerned).

Rather, for Glick, firms can preserve higher profit rates by preserving a cost advantage, and most pertinently by innovation involving more capital-intensive production techniques (though in principle this could also be through privileged access to cheaper inputs, including labour; the latter would also include the ability to differentially increase absolute rather than relative surplus value).

Type V^c statistics also relate to firms, and thus V_1^c , like V_1^b , measures the persistence of firm-level profits. However, recall Glick's assertion that the most desirable data would relate to line-of-business. All but the simplest (and smallest?) enterprises' products are in more than one line. From this point of view V^c statistics represent the lowest-available level of disaggregation, and are properly considered alongside the V^a statistics for the different SIC levels – representing an n -digit level of aggregation, where n is unknown (and inconsistent, in that larger firms produce a wider range of use values than smaller ones – see the FTC data in section 6.3.1 above). Thus V_1^c statistics would be comparable with V_1^a , not V_1^b .

6.3.2 Results

As pointed out in our review of Glick, his theory of competition maintains that profits are equalised by migration of capital to sectors with higher profit-rates from sectors with lower ones but provides no reason to exclude any particular sector of the economy from this

process. Thus there is no rationale for confining investigation to manufacturing industries alone, as is the case in his empirical investigation.

Our calculations of V_1^a are therefore made for two different sets of industries: first, for 18 manufacturing industries whose definitions (Table 6.4) are designed to correspond as closely as possible to those used in Glick's own study (remembering, as pointed out in the previous section, that these are not strictly speaking type V^a statistics); second, for all 61 industries defined at the two-digit level by the UK 1981 SIC.

Table 6.4: mapping from UK SIC codes to Glick manufacturing industries

Glick industry	1981 UK SIC industries
Food	41, 42 (excluding 429, tobacco, which is excluded by Glick)
Textile	43
Apparel	45
Furniture	467 (remainder of 46 excluded by Glick)
Paper	471, 472
Printing	475
Chemicals	25, 26
Petroleum	14
Rubber	48
Leather	44
Cement (stone, clay, glass)	24
Primary metals	22
Fabricated metals	31
Machinery	32, 33
Electrical equipment	34
Transport equipment	35, 36
Instruments	37
Miscellaneous	49

V_1 statistics for two-digit industries provide base-line results for work in the following sections. For each of the four profit rate measures examined Figure 6.2 compares V_1 statistics from Glick's own study (G) and the corresponding UK industries (aG), and V_1^a statistics for the whole UK economy (sic812).

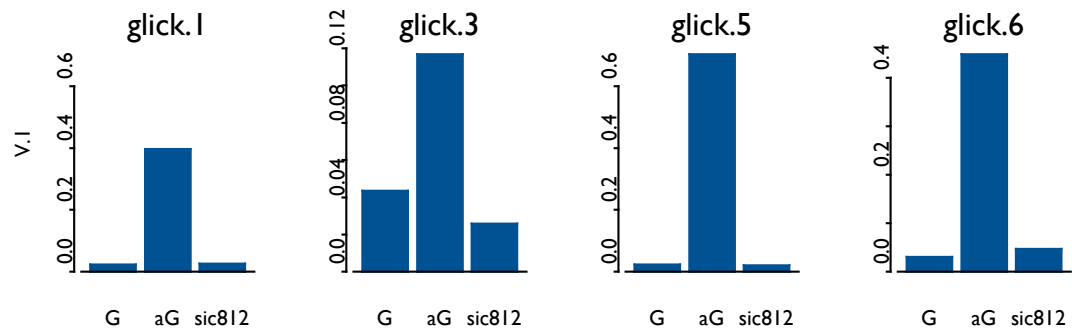


Figure 6.2: V_1 statistics for two-digit industries, Glick profit rate measures

As can be seen, the results imply a very much smaller degree of profit-rate gravitation within UK manufacturing than in the US, except in the case of Glick 3, though the improvement here is only apparent in relation to the other profit rate measures: the V_1 statistic is still more than twice as high as in Glick's own work. In contrast, the apparent gravitation across all two-digit industries is much greater, and on a par with that found within US manufacturing considered in isolation.

These results seem odd for two reasons. First, if dispersion within manufacturing is large, how can it be smaller for the whole economy, including manufacturing? Secondly, why should UK manufacturing show less evidence of gravitation than the US's? One might argue that this is explained by the exceptionally high concentration of industry in the UK, but on the other hand might this not be offset by the competitive pressure exerted by the UK economy's greater openness, measured by the value of trade relative to total output?

It is possible to outline hypotheses which might explain these results: the period in question was one during which UK exchange and interest rates experienced radical reversals (it includes the events of the so-called 'Black Wednesday' of 1992). Thus one might argue that the persistent within-manufacturing dispersion might be accounted for by the difference in fortunes between manufacturing sectors which were especially exposed to these events and those which were less so. (Recall that the reported V_1 statistics for manufacturing industry represent the variance of industry means around the *sectoral* mean, *not* about the mean for the economy as a whole).

However, while many would accept as a stylised fact that service industries in general are less exposed to overseas competition, it is also true that the UK non-manufacturing economy sector includes large financial and transport services industries as well as a significant primary industry with major overseas exposure (that is, oil and gas extraction). On this basis it is difficult to accept that gravitation in the non-manufacturing part of the economy should be significantly stronger. And one might also argue that those service sectors not exposed to the international economy might experience considerably less pressure towards equalisation.

A possible source of difference might be that the boundaries between two-digit UK SIC manufacturing industries are not the same as for Glick's classification (Table 6.4), but it seems hard to believe that this alone could account for such big differences. Somewhat more plausible would be to attribute the differences to the fact that the relative proportions of the different industries are different in the two economies.

We forego any attempt at further substantive investigation, as beyond the scope of the present work. For present purposes the important point is that we have demonstrated our claim that confining investigation of gravitation to the manufacturing sector alone, rather than to the economy as whole as demanded by the classical theory of competition, may indeed be seriously misleading.

6.4 Testing Glick

6.4.1 The level of aggregation

In this section we test to see whether varying the level of aggregation at which Glick's V_1 statistic is calculated produces the expected results. Defining the movement from the four- to the one-digit level of SIC classification as an increase in level of aggregation, and a movement from higher to lower values of V_1 as an increase in degree of gravitation, we should expect a positive correlation (see section 6.3.1).

In our review of Glick we noted that he himself suggests that tests of gravitation using data aggregated to the two-digit industry level may well be inappropriate; the particular

notion of ‘market’ to which the classical theory of competition refers is most closely approximated by the ‘line of business’ data briefly collected in the US. From this perspective even firm-level data, such as that available through FAME, is potentially at too high a level of aggregation, especially in the case of larger firms.

We apply his methods to our FAME dataset once again, but now also to industries aggregated at the one-, three- and four-digit levels of the 1982 UK SIC system. We also calculate V_1 for all *firms*; this may be considered as a test at an n -digit level, where n is an unknown, and possibly varying, number greater than four; we therefore expect to find less evidence of profit-rate gravitation than in the other four cases. As in the previous section we restrict ourselves for the time being to the four best Glick profit rate measures.

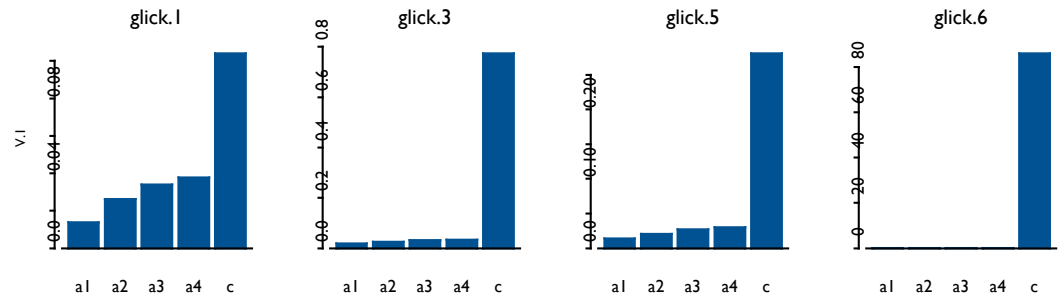


Figure 6.3: V_1 statistics for different levels of aggregation, Glick profit rate measures

In Figure 6.3 bars labelled ‘a1’ to ‘a4’ represent values of V_1^a for one- to four-digit industries, and bars labelled ‘c’ represent V_1^c calculated across *firms*, not industries (note that the vertical scales vary from panel to panel).

The striking result to be seen from the Figure is that the aggregation/gravitation correlation is apparently strongly positive, as expected: V_1 is smaller, hence apparent profit-rate gravitation higher, at lower levels of aggregation, including at the notional n -digit level constituted by firms as a further level of aggregation below the four-digit one (sufficiently pronounced, in the case of Glick 6, to mask the results in respect of the other levels).

6.4.2 The measure of dispersion

Glick's preferred measure of profit-rate gravitation is V_1 , the variance of the time-averages of industry profit rates. We have criticised this on the grounds that it does not tell us how to compare two cases where V_1 is equal but V , the overall variance, is unequal. To meet this difficulty we have suggested that a natural alternative is the ratio $\hat{V} = V_1/V$, in other words the *proportion* of V which is accounted for by V_1 . The possible significance of this can be seen by referring to Figure 6.4 and Figure 6.5 (here the panels have common vertical scales.)

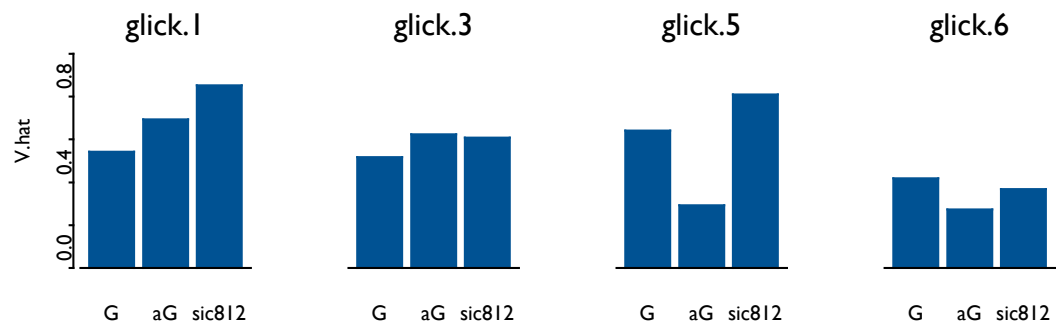


Figure 6.4: \hat{V} statistics for two-digit industries, Glick profit rate measures

Figure 6.4 corresponds to Figure 6.2 above, in presenting a comparison of the results from Glick's own work with equivalent data from the UK economy. Using \hat{V} as the measure of gravitation substantially alters the previous results: where we saw much less gravitation in UK manufacturing than in the US, we find it comparable or greater, while whole-economy gravitation is now greater than or comparable to that within the manufacturing sector.

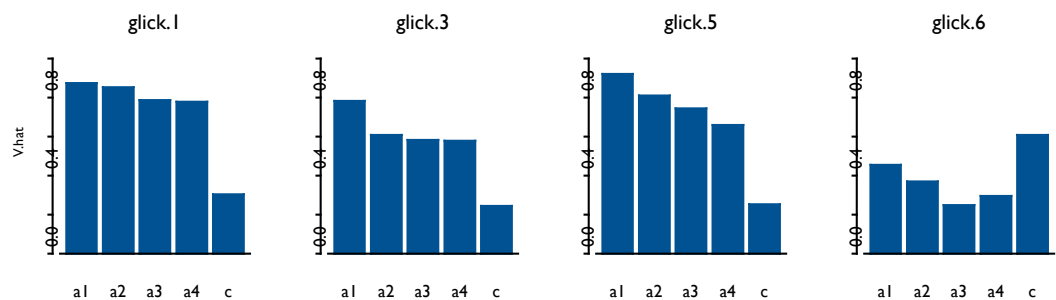


Figure 6.5: \hat{V} statistics for different levels of aggregation, Glick profit rate measures

Figure 6.5 compares with Figure 6.3 above, showing the effect of testing at different levels of aggregation. In this case the qualitative picture is *reversed*, resulting in a perverse (that is, negative) correlation of gravitation with aggregation (including the results in respect of \hat{V}^c , except in the case of Glick 6).

Use of \hat{V} , therefore, allows us to obtain results broadly consistent with Glick's original work, in terms of two-digit industries; however, it undermines the apparent confirmation of our expectations about the relationship between aggregation and gravitation.

6.4.3 Alternative profit rate measures

It is pertinent to ask how use of this modified test statistic might alter conclusions about the appropriate profit rate measure for testing gravitation. We do this by comparing results for 12 profit rate measures, first using V_1 and then using \hat{V} . The profit rate measures are the four Glick measures already investigated together with the four standard accounting ratios and the four 'traditional marxist' measures identified by Gillman (discussed in Chapter Two).

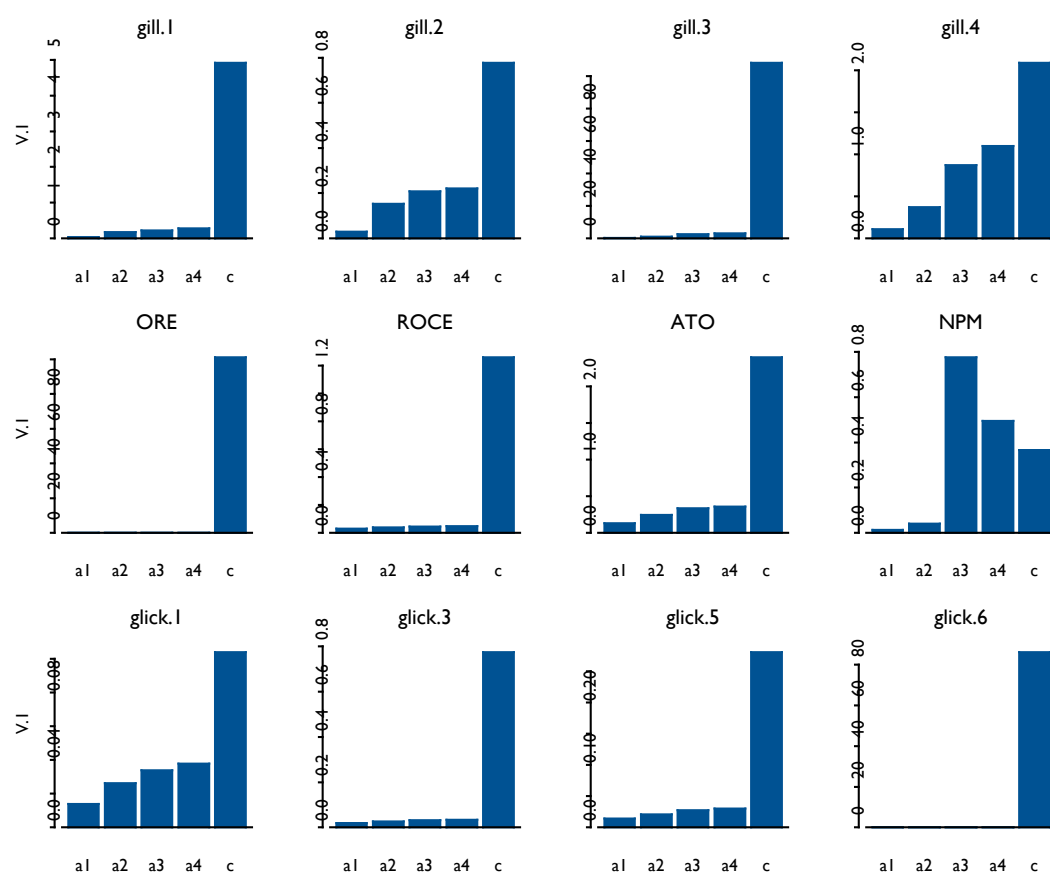


Figure 6.6: V_1 statistics for different levels of aggregation, multiple profit rate measures

Figure 6.6 shows the results with Glick's preferred V_1 statistic, including those previously obtained for the Glick measures.⁷⁶ As can be seen, the expected pattern of correlation between aggregation and gravitation obtains for the eight extra measures (with NPM, net profit margin, being idiosyncratic). Also repeated is the consistency of the results for firm-level gravitation with the view that this is simply a further level of aggregation below the four-digit one.

But the most important result is that Gillman's 'marxist' measures do very much worse than any of Glick's measures, in the sense of having much larger V_1 statistics, at all levels of aggregation (with the exception of the V_1^c statistics for Glick 3 and Glick 6).

⁷⁶ See Figure 6.3 above: as there the vertical scale varies from panel to panel.

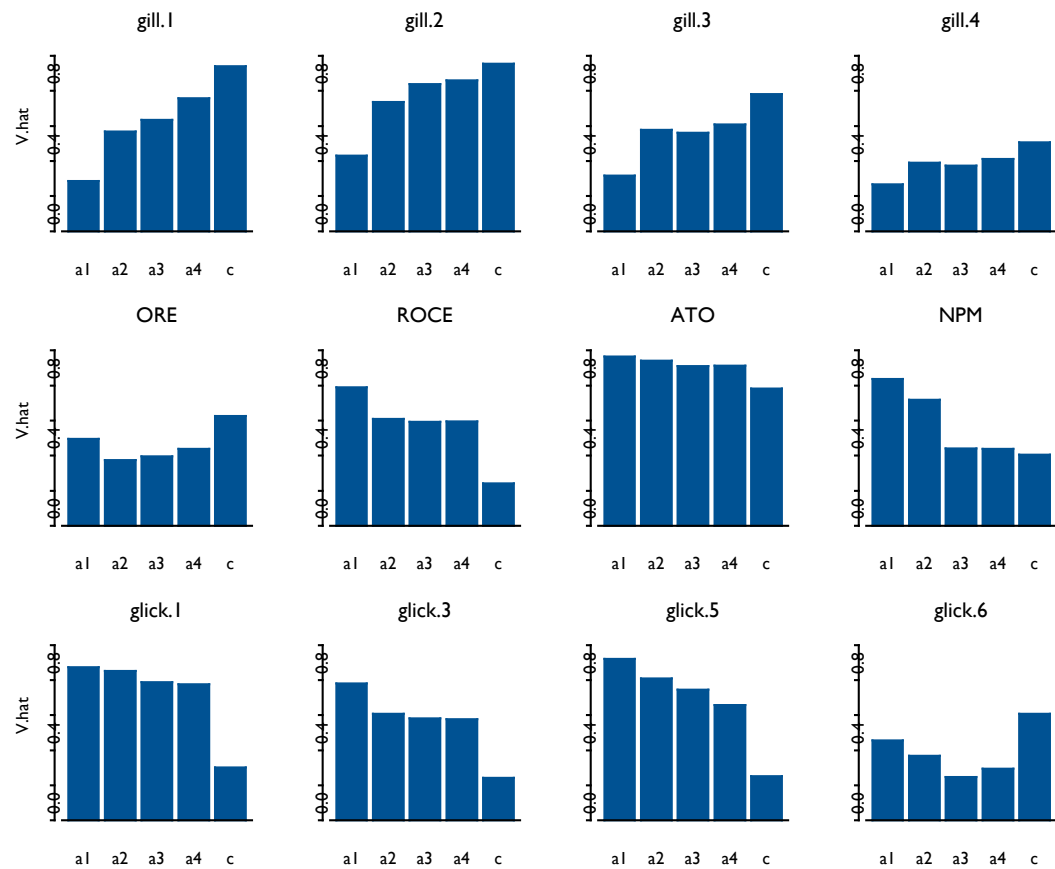


Figure 6.7: \hat{V} statistics for different levels of aggregation, multiple profit rate measures

Figure 6.7 shows the results with \hat{V} , again including those previously obtained for the Glick measures.⁷⁷ Once again, considering the value of V_1 relative to the total variance V radically alters the conclusions to be drawn.

Above, we saw that for Glick's preferred profit rate measures the apparent correlation between aggregation and gravitation became perverse when considering \hat{V} . This perverse correlation also appears in the case of three of the accounting measures (where ORE and its analogue Glick 5 are now idiosyncratic).

⁷⁷ See Figure 6.5 above; as there, the panels have a common vertical scale.

But in the case of the Gillman measures this perverse correlation is not merely diminished, but reversed in sign. Apparent profit-rate gravitation is now increased by aggregation, as one would expect.

Moreover, the supposed greater evidence of profit-rate gravitation provided by Glick's profit rate measures disappears. The relative variance of his best measures is no smaller than that of any of the Gillman measures, indicating that they find the same or less evidence for gravitation. Indeed Gillman 4 is now better than any other measure bar Glick 6, which like its fellows suffers from the perverse correlation complained of above.

6.4.4 Conclusion

Our internal critique of Glick focussed on his use of data for manufacturing industries alone, rather than the whole economy, his choice of profit rate measure, the level of aggregation of the data, and his chosen test statistic. Replicating and extending his methodology using our U.K. data set has produced three results which supported all four aspects of this critique.

- His test statistic produces an unlikely relation between the evidence for gravitation in two-digit manufacturing industries and two-digit industries for the economy as whole
- Using the relative measure of gravitation $\hat{V} = V_1/V$ corrects this, at the expense of showing a perverse relation between aggregation and gravitation
- Using Gillman's Marxian profit rate measures and the relative measure of gravitation not only produces the expected relation between aggregation and gravitation, but also results in better evidence for gravitation than Glick's preferred measures.

6.5 Industry distributions: the *L*-moments perspective

The preceding sections of this chapter have established that Glick's work on profit gravitation is indeed vulnerable to the 'internal' critique levelled against it in both Chapter Two (section 2.2) and the current chapter (section 6.2). This section will implement our

‘external’ critique of Glick by applying our technique of using random-sized random samples to estimate L -moments.

Our external critique included the points made in sections 6.3.3 and 6.3.4. The first point was that he does not test the full implications of his hypothesis, in that he does not show that *intra*-industry profit rate dispersions are indeed bigger or more persistent than *inter*-industry ones, as his discussion of dispersion implies. The second was that Glick does not look systematically at higher moments of profit rate distributions.

We suggested that this thought might inspire two possible strategies: first, follow the line implied by Glick’s discussion, and investigate the higher moments of industry average profit rates; second, investigate whether different industries appear to follow different distributional models.

If we examine the means and co-efficients of L -variation, as well as the L -skewness and L -kurtosis, this second strategy can also shed light on the question of the scale and persistence of intra-industry dispersion; therefore it is the one we shall pursue.

We thus intend here to use our perspective and methods to test both the general robustness of Glick’s theoretical perspective and the appropriateness of his preferred profit-rate measures in comparison with alternatives.

6.5.1 Methods

Objectives As carried out in Chapter Five, we will estimate the L -statistics and plot the results in L -skewness/ L -kurtosis space, this time for each industry to be considered, for each of the years for which data is available. This will provide direct evidence on the comparability of different industry distributions, and on their persistence through time. However, we will lay more stress than before on the use of plots in L -moment ratio space – in particular the coefficient of L -variation (cLv)/ L -skewness and mean/ cLv planes – to identify relationships between skewness and dispersion.

We wish to further restrict the scope of the work, in order to avoid both an excessively diffuse discussion and a potentially overwhelming burden of computation; we do this by further selection within the set of profit rate measures and by confining the investigation to a single level of aggregation. Our primary focus will be on the two-digit industries defined in the 1981 SIC and on five profit rate measures, using the same data sets as used in the previous sections of this chapter. The profit rate measures will be Gillman 3 and Gillman 4, Glick 1 and Glick 5, and the asset turnover rate (ATO). We justify these choices as follows.

With regard to the profit rate measures, we have shown that our relativised version of Glick's test statistic avoids problems in reproducing Glick's work with UK data, at the expense of destroying the ability of his preferred profitability measures to exhibit the theoretically-expected pattern of variation across different levels of aggregation. In contrast, not only do all the Gillman measures examined above demonstrate the expected relationship, but the two selected here demonstrate greater evidence of profit-rate gravitation than any of the Glick measures. Among the latter we choose the two measures which perform best in Glick's own work. In short, we are selecting the two measures most preferred by ourselves and the two most preferred by Glick.

ATO is not chosen because of any suggestion that it is subject to gravitation. Rather, we select it because we know from the empirical density function that the whole-economy distribution across firms shows a unique bi-modal structure (Figure 3.5), which it is easy to show is attributable to clear differences of distributional form within one-digit SIC industry 8, financial services (see below). Whether this bi-modality is carried through into the distribution of rates of return at the capital level is harder to judge. Certainly there is some suggestion of it in the empirical density function (Figure 4.6), but on the other hand this plot also has some of the features of an exponential distribution, an impression reinforced by the *L*-moment ratio analysis of Chapter Five (Figure 5.3). This measure thus provides a clear demonstration of the potential of the methods employed here.

The choice of the two-digit level is taken in the interests of comparability with Glick's own study, although our work in section 6.5 strongly suggests that a wider study should look at the internal distributions at the four-digit level.

Data For strict comparability of results within the present chapter we further choose to use the same FAME company data sets used in the earlier sections of this chapter; that is, ones for each profit rate measure comprising all those firms for which the relevant measure can be computed for all five years of available data.

However, we note that for a broader study of this question this is not necessarily the most desirable choice. Our own view is that any theory of profit rate distributions must apply to the whole population of firms existing at any particular moment, and not to a subset composed of those which happen to survive some particular time interval, especially if that time interval is chosen arbitrarily on the basis of what data happen to be available to the investigator.

A particular consequence of our choice of data sets is the disadvantage is that the results presented below are not comparable with those of Chapter Five. The decision to restrict ourselves to sets of surviving firms has the further consequence that for some industries and some profit rate measures the number of firms concerned is too small to compute the full set of L -moments, with the result that we have to exclude these cases. However, although widening the study to include all firms in each year would tend to raise the number of firms included in each industry, working at the preferred four-digit level would tend to lower it again.

Sampling The objectives outlined here imply a substantial computational effort: at the two-digit level there are up to 61 industries to be examined for each profit rate measure tested, and we have five years' data for each industry. The RS2 technique used in previous chapter involved taking 100 samples for each profit rate measure in each year, and each of these was constructed by concatenating 100 sub-samples (to cope with the fact that the extreme ranges of the capital measures could otherwise lead to samples to be too small to

be of use). We then have to calculate the first four L -moments for each sample, and the time taken for this increases rapidly with size of sample.

The present case involves working with up to $5 \times 61 \times 5 \times 100 = 152,500$ samples; since within-industry dispersion of firm size is likely to be smaller than the overall dispersion, we can take advantage of this to reduce the number of sub-samples. The ideal method would involve making the number of sub-samples depend on the shape of the firm-size distribution, but this is itself a complex topic. We therefore take the simpler approach of setting the number of sub-samples at $500/n$, where n is the number of firms in the industry.

6.5.2 Results

The L -moment analysis of inter-industry variation yields three main results, demonstrating (1) why Glick's preferred measure of gravitation performs so differently from our relativised version $\hat{V} = V_1/V$, (2) that skewness of within-industry distributions does indeed undermine the validity of inter-industry comparisons relying on means, and (3) that L -moments distinguish distributional models in relation to different industries as well as to different profit rate measures.

Relative versus absolute gravitation Glick seeks to demonstrate the existence of classical gravitation by showing that the total variance of the time-averages of industry profit rate deviations is small, a quantity he denotes by V_1 . As we have shown above this measure has the defect that it fails to provide a conclusive criterion in cases where V_1 is equal but total V (the total variance of annual industry rates) is unequal.

This can be removed by considering instead the relativised measure $\hat{V} = V_1/V$, but when applied at different levels of aggregation this produces counter-intuitive results for his preferred measures of the profit rate. These counter-intuitive results do not apply in the case of those measures which we regard as the appropriate marxist measures, and moreover these latter display more evidence of gravitation than do Glick's at the levels of aggregation

lower than the two-digit one (used by Glick, but admitted by him to be an excessively high one).

The following charts indicate why this is so. For each of the five profit rate measures to be considered in this section, Figures 6.8 to 6.12 plot annual industry mean profit rates against annual co-efficients of L -variation (cLv) for each of the two-digit industries for which the statistics can be estimated by our RS2 procedure.⁷⁸ The plotting characters denote the one-digit sector into which each two-digit industry falls; thus in a sector with five two-digit industries we will see plotted up to 25 observations (five industries in each of five years).

To provide comparability without obscuring differences in the structure of the data, the scales for the mean are constrained within the range -0.25 to 2.0 with the actual minimum (maximum) being the minimum (maximum) of the data if greater (less) than the standard limit; the scale for the cLv is from 0 to 1.0 in all cases.

⁷⁸ We are using a much-reduced data set compared to that used in Chapters 3 to 5; the restriction to firms for which there is data for all five years, coupled with the extreme skewness of firm size, means that even at the two-digit level there are in some industries too few firms to allow calculation of sample L -moments.

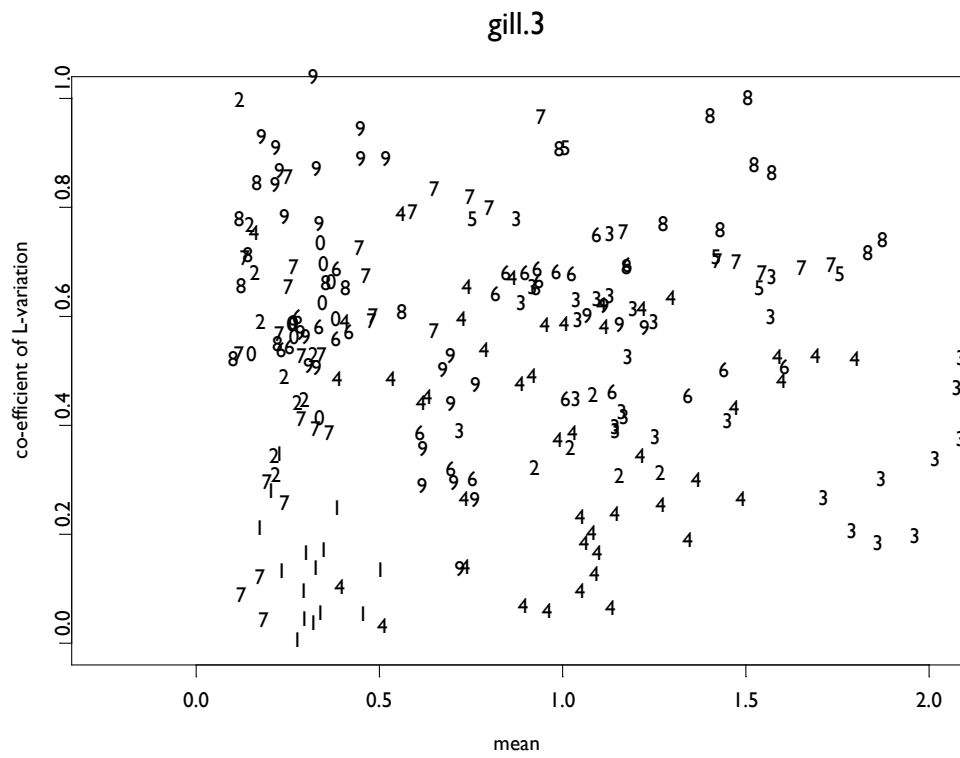


Figure 6.8: Gillman 3, mean/cLv space

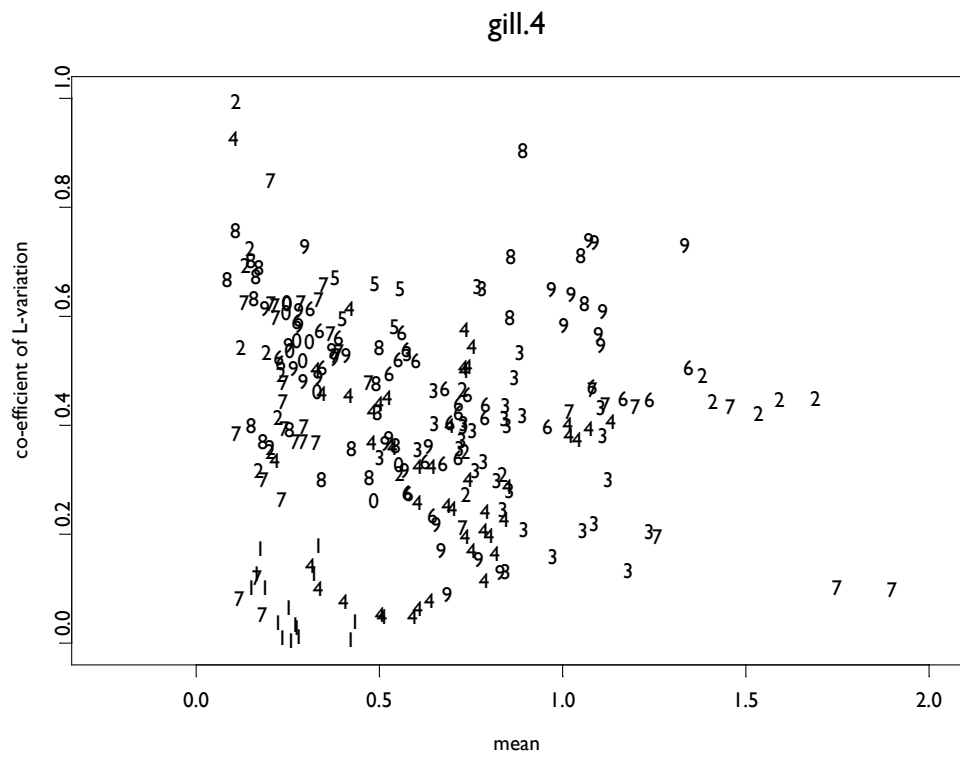


Figure 6.9: Gillman 4, mean/cLv space

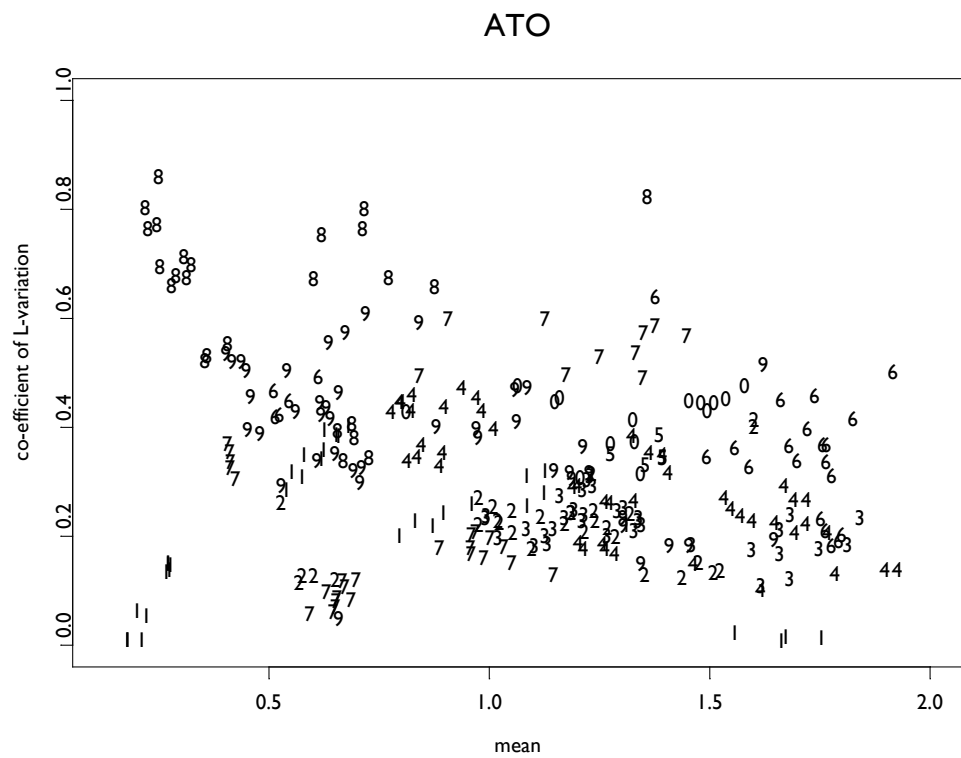


Figure 6.10: ATO, mean/cLv space

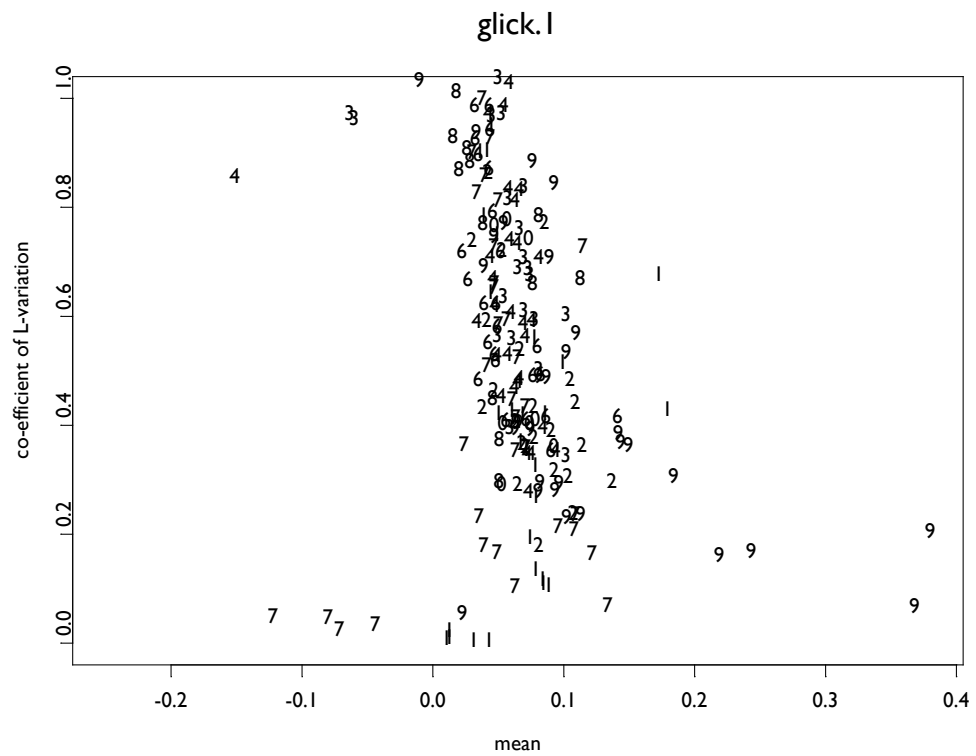


Figure 6.11: Glick I, mean/cLv space

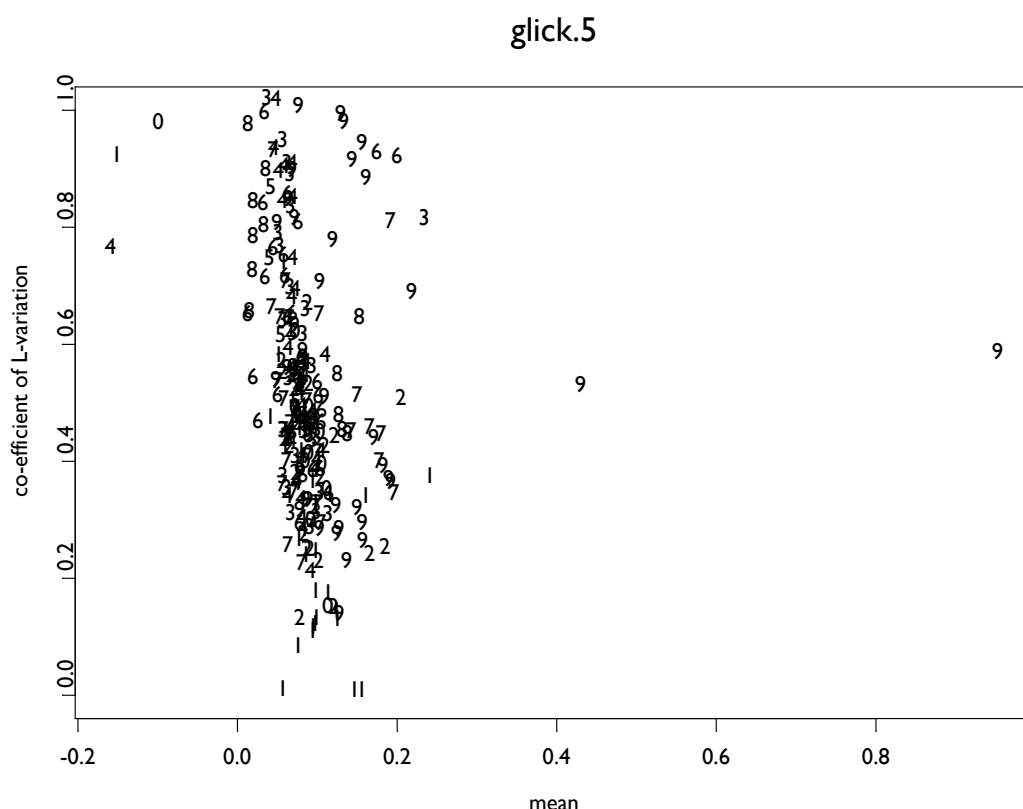


Figure 6.12: Glick 5, mean/cLv space

As can be seen from the means for the Gillman 3 and 4 measures and ATO (Figures 6.8, 6.9 and 6.10 respectively) are (very roughly) evenly distributed over the range between zero and 2.0. By contrast, the means of Glick 1 and 5 are highly concentrated around the 0.1 mark (Figures 6.11 and 6.12 respectively).

Recall that at the two-digit level the profit rate measure Gillman 4 had a \hat{V} score of 0.4, whereas Glick 1 and Glick 5 had values in excess of 0.8, Glick 3 one of 0.6, and only Glick 6 improved (marginally) on 0.4. The implication is that notwithstanding the lower V_1 statistics (the variation of the industry time-averages) of the Glick measures when compared to other measures, they are large compared to the overall variation V of industry profit rates because the latter is also very low.

Implications of within-industry skewness We have previously pointed out (section 6.3.4) that if within-industry profit-rate distributions are skewed the mean is not obviously better than, say, the mode as a statistic for assessing the gravitation of industry

profit rates. The most extreme case is that of two industries with distributions which are similar in shape except for having opposite skewness; in this case equality of means implies unequal modes and vice versa, and in either of these two eventualities a mean-based test of gravitation will give the opposite answer to a mode-based test.

The next set of figures plots RS2 estimates of L -skewness and L -kurtosis for five profit rate measures. As in the previous section each plot shows five years of observations for two-digit industries, identified by the one-digit sector to which they belong. In each plot we show the loci of a number of three-parameter location-scale-shape distributions given by Hosking and Wallis (1997); as in Figure 4.1 they are the generalised logistic (GLO); the generalised extreme value (GEV); the three-parameter lognormal (LN3); the three-parameter gamma, or Pearson Type III (PE3); the Weibull; and the generalised Pareto (GPA), considered from top to bottom at L -skewness = 0.25; also shown are the loci of several two-parameter location-scale distributions: the uniform (■ in the Figure), the Gaussian (◆), the exponential (●), and the Gumbel (▲). Also as in Figure 4.1 the shaded area indicates combinations of L -skewness and L -kurtosis which no distribution can have.

At first glance the most obvious feature (Figures 6.13, 6.14 and 6.15 respectively) is that the two Gillman measures and ATO all display a marked tendency to positive skewness (in the case of Gillman 3, of a fairly extreme kind), whereas within-industry skewness of the two Glick measures is clustered symmetrically about zero (Figures 6.16 and 6.17 respectively).

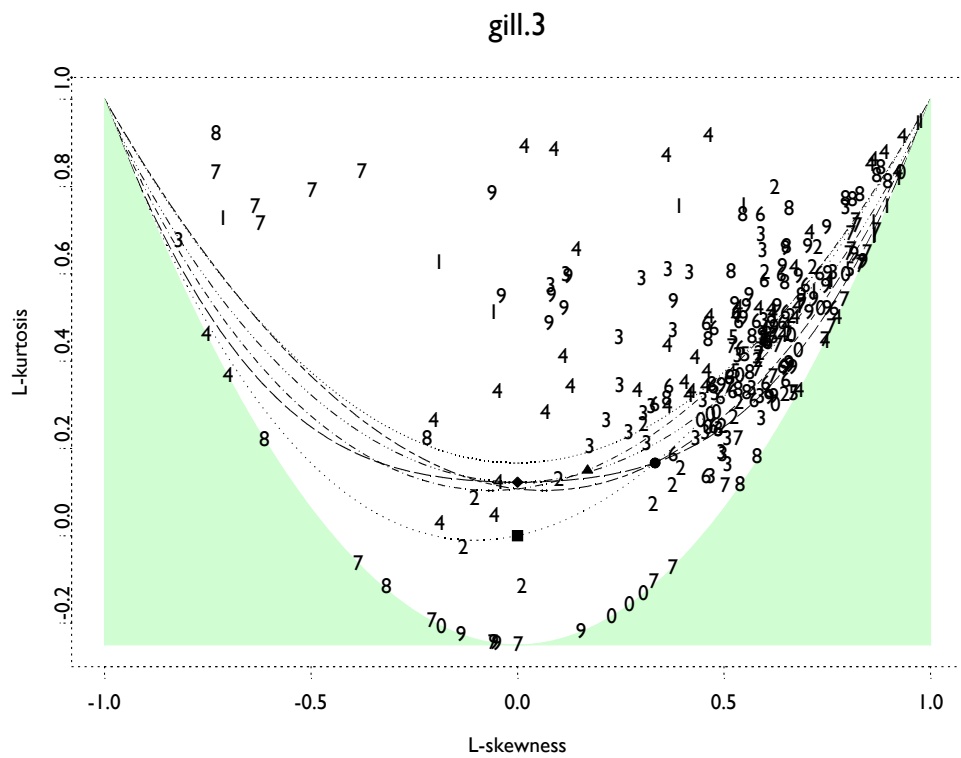


Figure 6.13: Gillman 3, L-skewness/L-kurtosis space

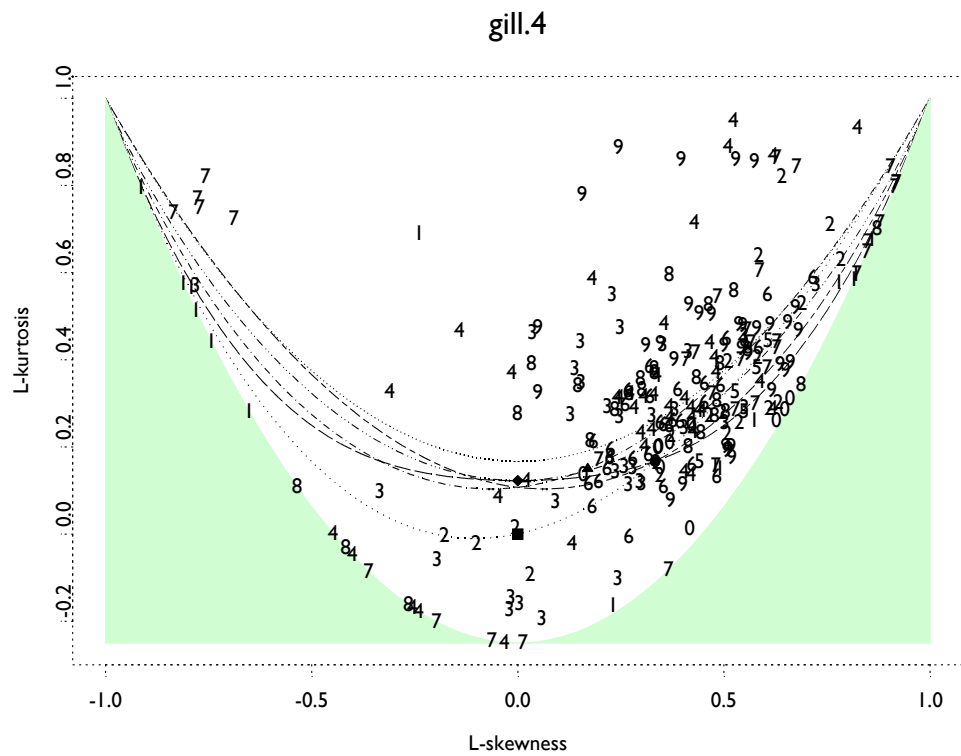


Figure 6.14: Gillman 4, L-skewness/L-kurtosis space

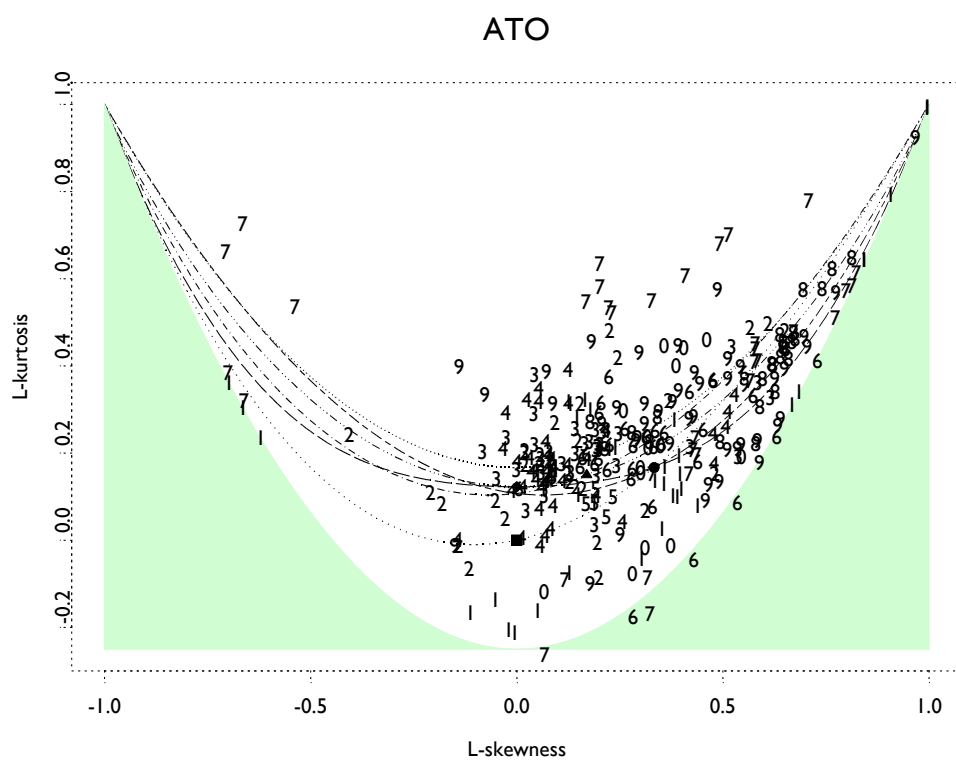


Figure 6.15: asset turnover (ATO), L-skewness/L-kurtosis space

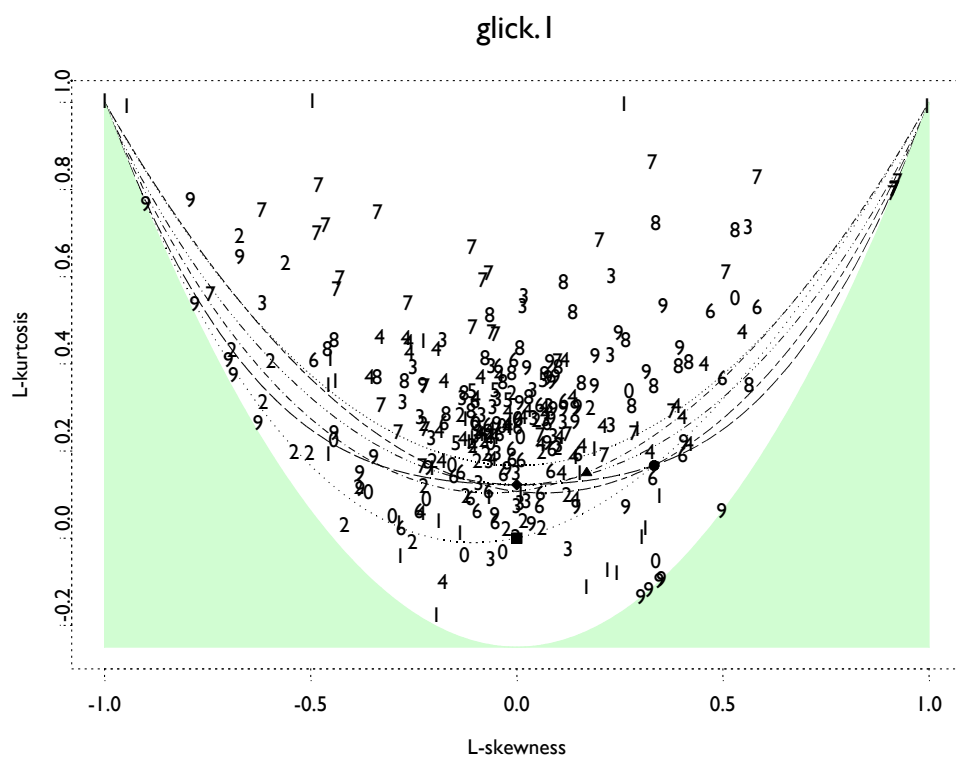


Figure 6.16: Glick 1, L-skewness/L-kurtosis space

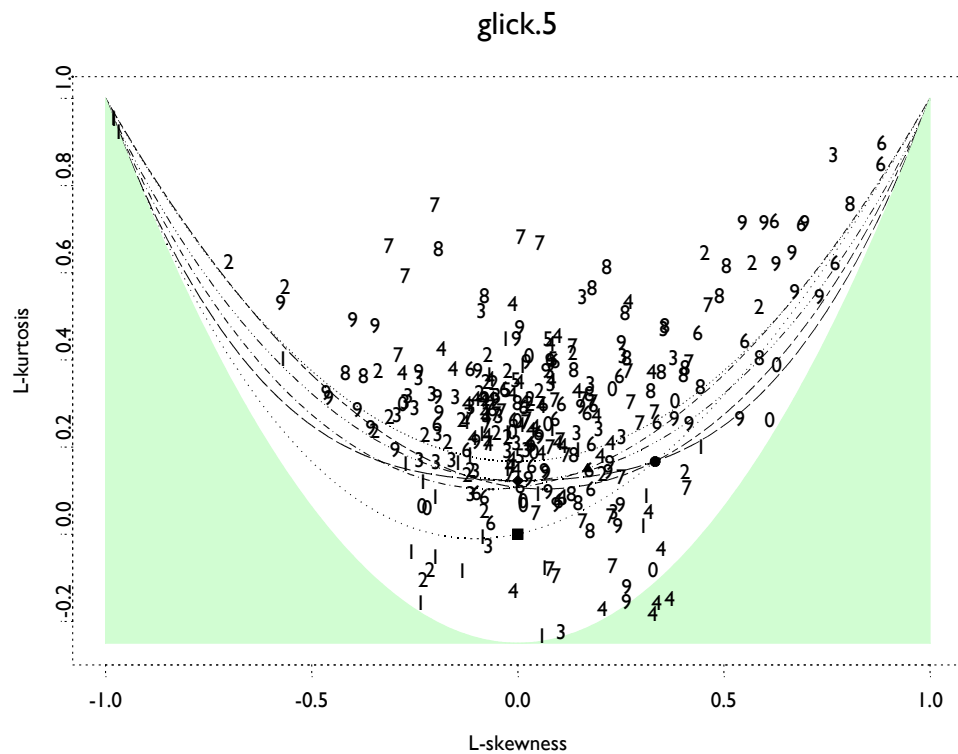


Figure 6.17: Glick 5, L -skewness/ L -kurtosis space

It might be thought that this implies that any problem caused by skewness is thus less severe for Glick's preferred measures, but in fact the reverse is the case: taking location in L -skewness/ L -kurtosis space as a summary of the shape of a (unimodal) distribution, what figures 6.13 to 6.17 show is that for any arbitrarily-chosen industry there is a high probability of finding another industry the distribution of which has similar shape but opposite skewness, precisely the extreme case envisaged above.

Moreover, as seen from Figures 6.18 and 6.19, for Glick 1 and 5 the distribution of industries in cLv/L -skewness space is also approximately symmetrical, implying that for an industry whose distribution has arbitrary scale there is likely to be one of similar scale but opposite skewness.

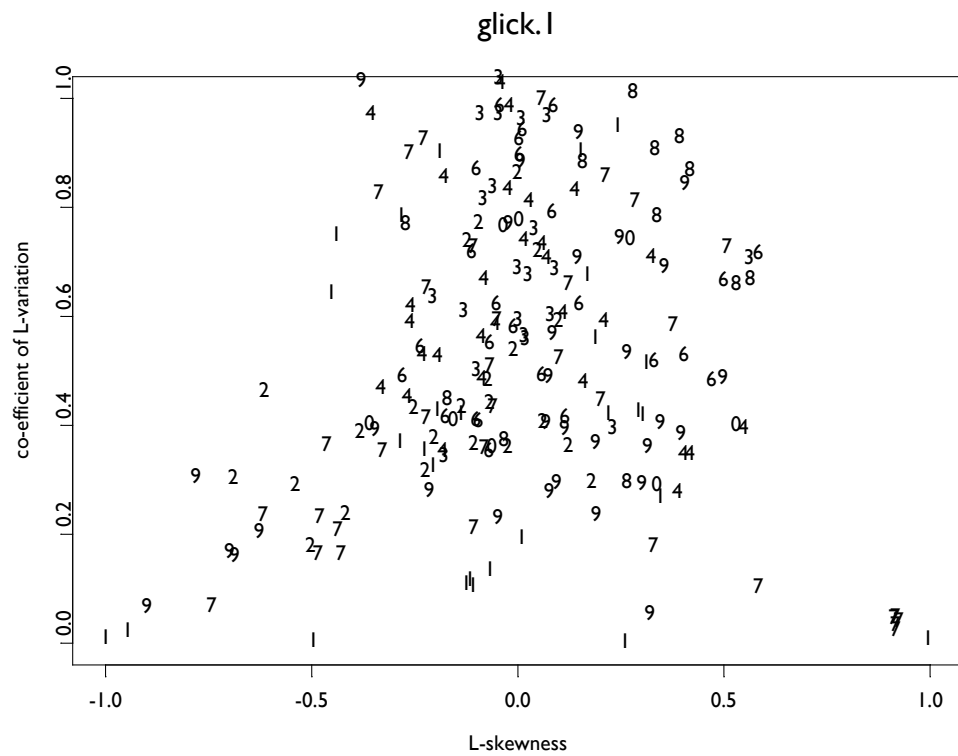


Figure 6.18: Glick 1, L-skewness/cLv space

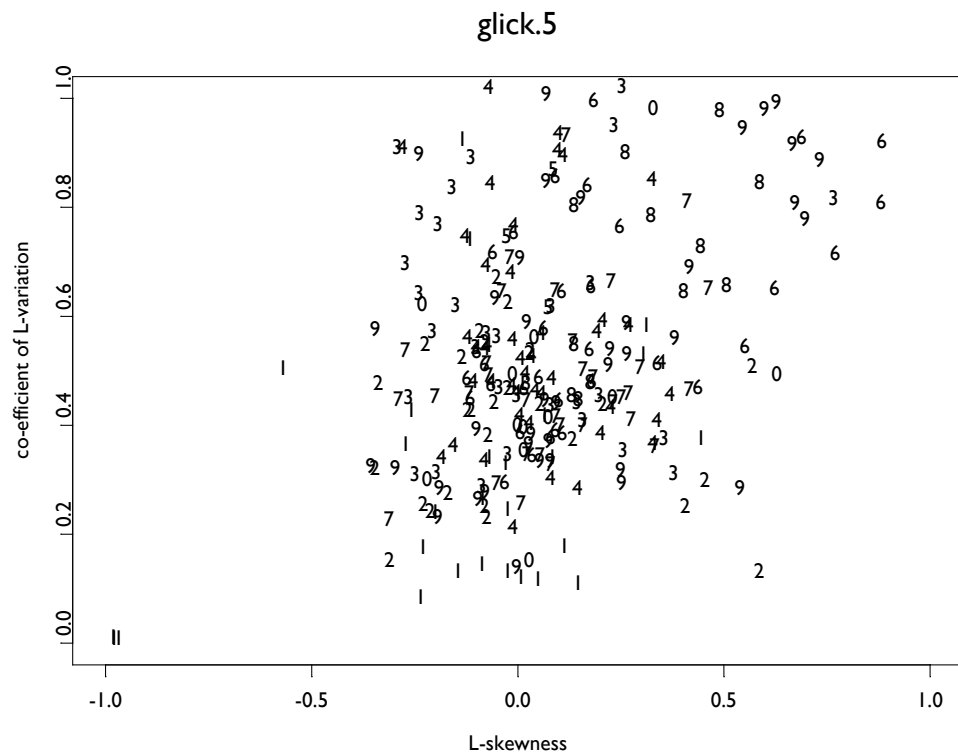


Figure 6.19: Glick 5, L-skewness/cLv space

Taken together these facts suggest that, if one measures two-digit industry profit rates according to Glick's preferred measures, the skewness problem suggested here is likely to be prevalent: for any particular industry there is likely to be another whose distribution is a mirror image of the first – and if their means exhibit any tendency to gravitation their modes will tend to diverge. Thus it should not be a surprise if means-based tests produce inconsistent or counter-intuitive results.

In contrast, the other profit rate measures examined here show not only strong positive correlation between skewness and kurtosis, but (not shown here) also some evidence for positive correlation between skewness and scale; thus the mirror-image case should be rare, and to that extent so should that of mean- and mode-based measures conflicting.

6.5.3 Resolving power of *L*-moments

Of the total of 21 profit rate measures examined at one point or another of this work, the only one to exhibit any trace of bi-modality over the economy as a whole is ATO. Here we show that this overall bi-modality results from a mixture of different distributions relating to different industries, and that these different models are successfully distinguished by *L*-moment analysis.

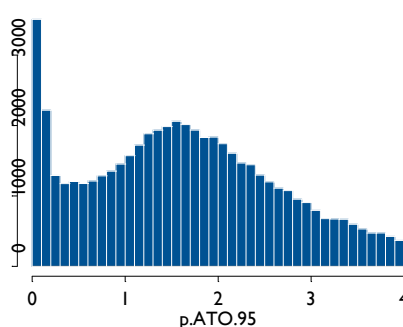


Figure 6.20: asset turnover rate (ATO), whole economy, firm-level distribution

Figure 6.20 is an enlarged version of the relevant panel of Figure 3.5, the empirical density distribution at the firm level estimated by histogram. Figure 6.21 shows histograms of ATO for each one-digit sector; as can be seen, the sectoral distributions vary considerably in form. Note the top left panel: this shows sector 8, banking and finance, to

have a pronounced spike at approximately 0–25 per cent (as does sector 7, extreme right of middle row).

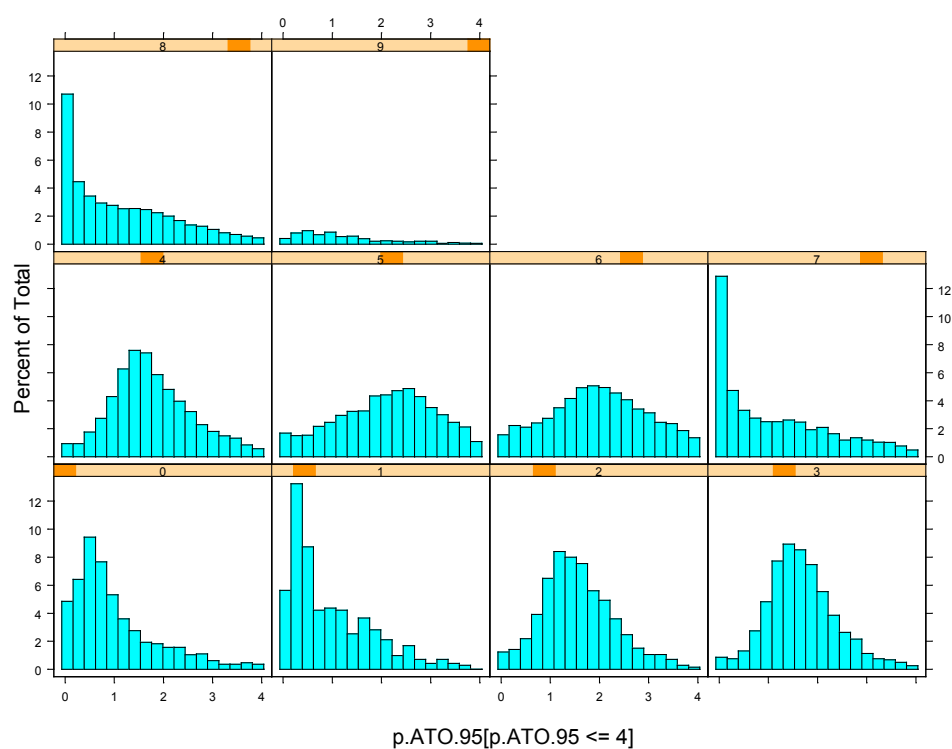


Figure 6.21: ATO firm-level distribution by one-digit sector

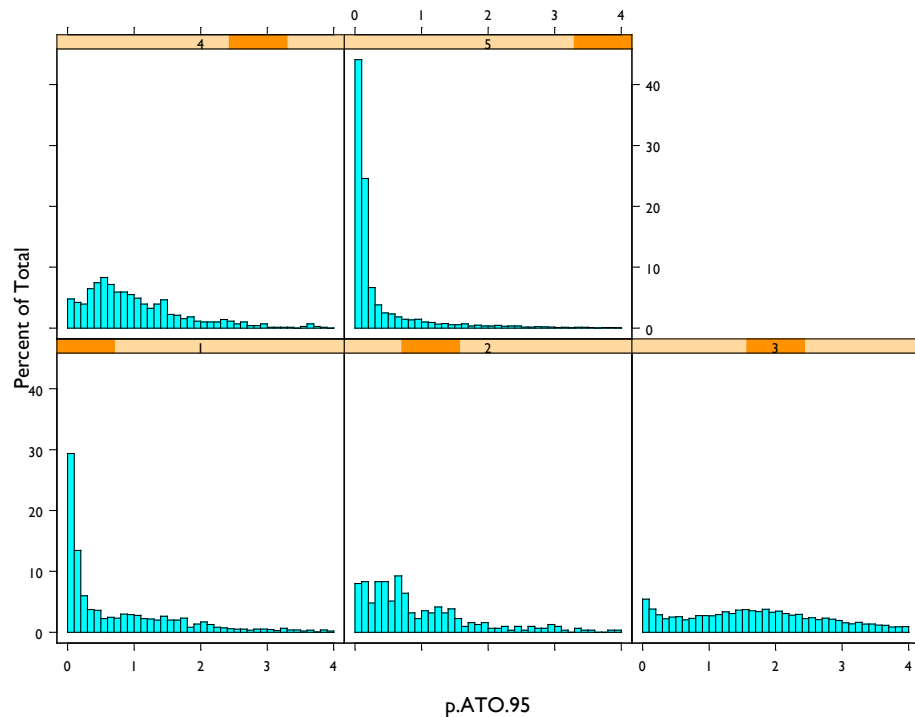


Figure 6.22: ATO firm-level distribution; sector 8 by two-digit industry

Moving to the two-digit level, the histogram density estimates for industries 81 to 85 are given in Figure 6.22. We see that the left-hand spike in sector 8 is largely due to the distributions of industries 81 (banking and finance: bottom left panel) and 85 (owning and dealing in real estate: right hand of upper row).

Of the three, the biggest contribution comes from industry 85, which is larger than 81. It will be noted that industry 83 exhibits strong bimodality on its own; investigation of this at lower levels of aggregation seems likely to reveal a further pattern of mixtures of distributions.

The charts above are of the profit rate distribution across firms, irrespective of their size, as measured by the relevant capital measure. We now show that qualitatively similar results flow from applying our RS2 method of randomly-sized capital-weighted samples.

Figures 6.23 to 6.25 show three cross-sections of L -moment ratio space – respectively, the L -skewness/ L -kurtosis, L -skewness/ cLv and mean/ cLv planes – with five annual

estimates for each of the two-digit industries within sector 8, plotted using the appropriate second digit. These reveal that with the exception of industry 84 (renting of movables) all industries in the sector have exceptionally high skewness, kurtosis and coefficients of L -variation (cLv). They also suggest that the concentration of rates around 0.1 is largely due to industries 81 (banking and finance as such) and 85 (owning and dealing in real estate), that is, industries whose ‘profits’ are in fact rents. All this is consistent with the evidence of the histograms.

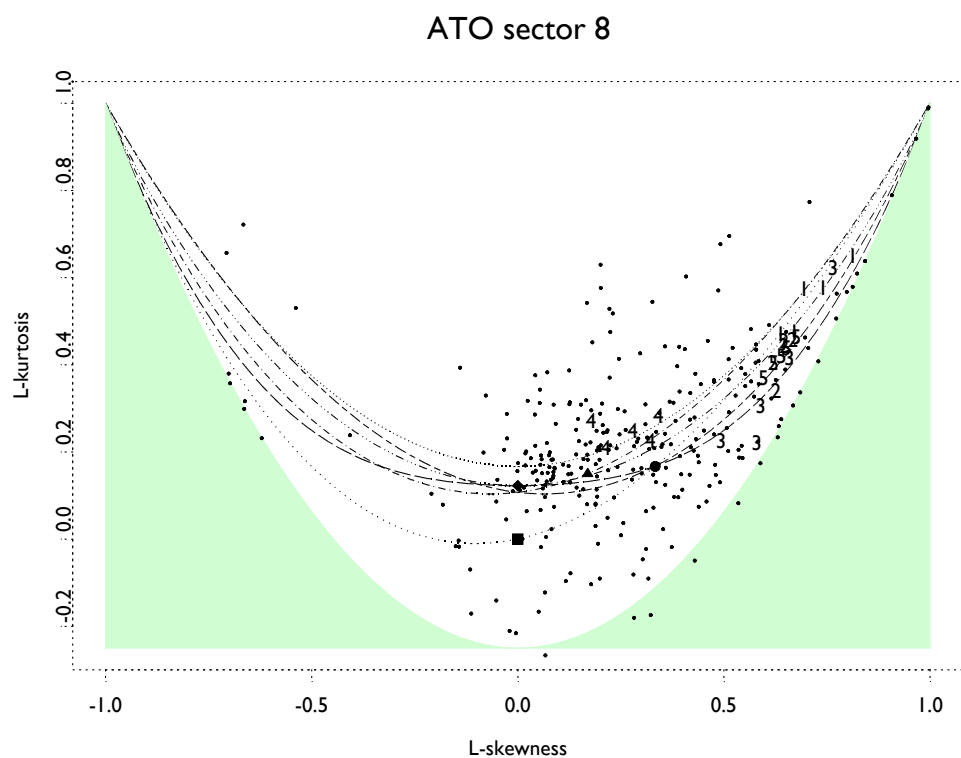


Figure 6.23: ATO, L -skewness/ L -kurtosis space, sector 8

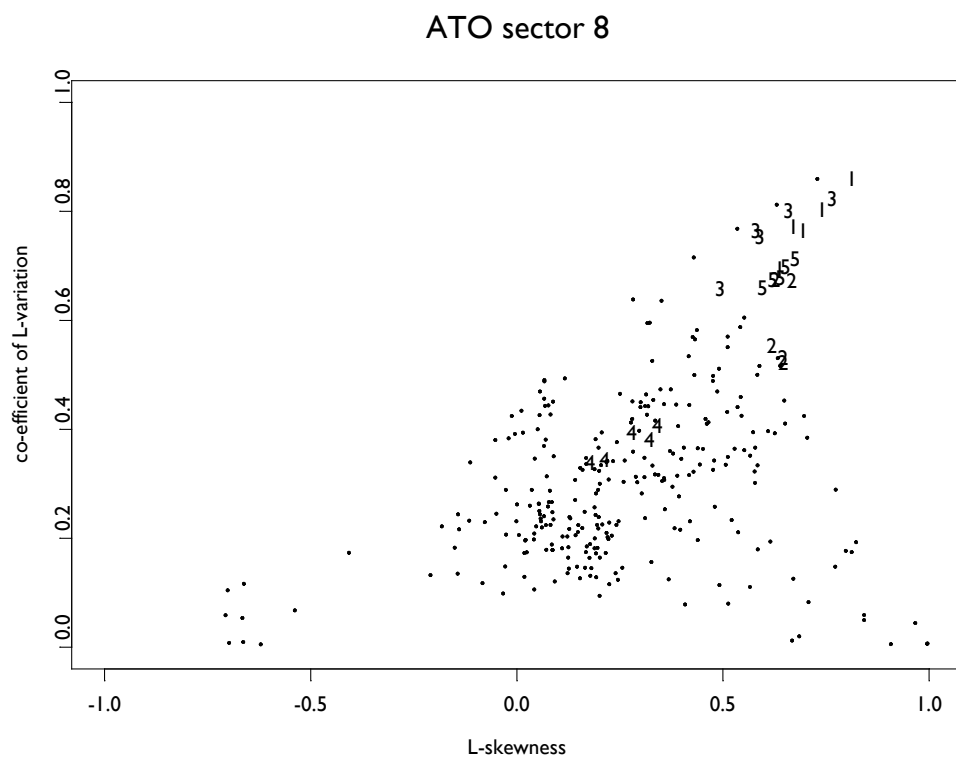


Figure 6.24: ATO, L-skewness/cLv space, industries 81-85

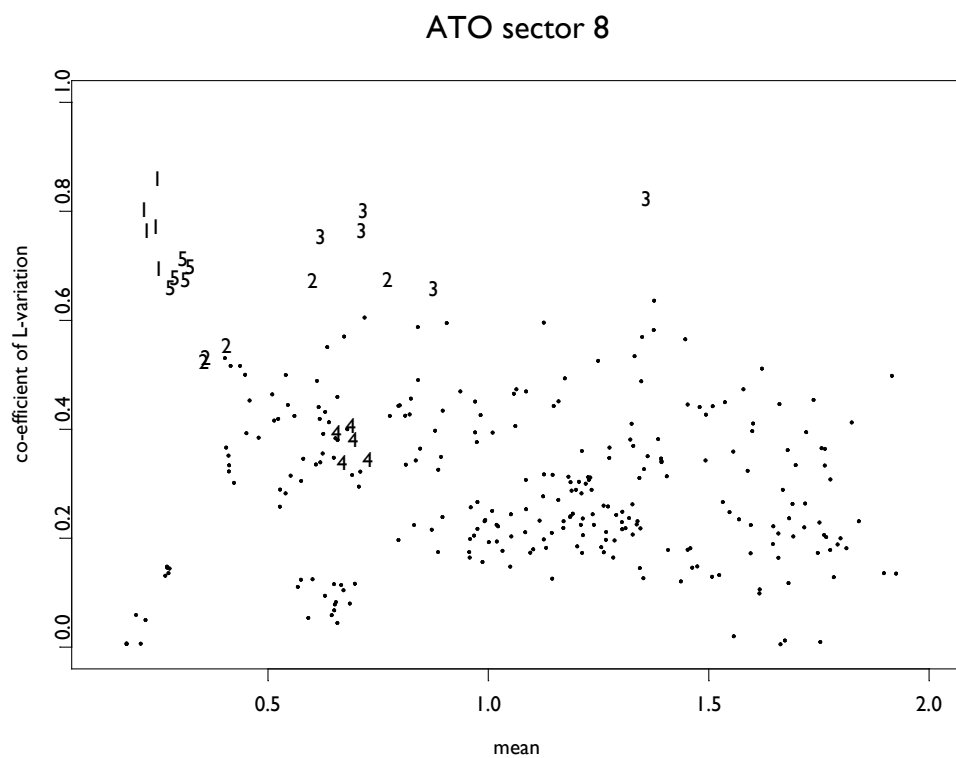


Figure 6.25: ATO, mean/cLv space, industries 81-85

6.6 Conclusion

We have argued that Glick's (1985) tests of the gravitation of industry profit rates may be questioned on several theoretical and methodological grounds: that they are at an excessively high level of aggregation (as suggested by Glick himself), are wrongly confined to consideration of manufacturing industry alone, use a questionable test statistic, focus on inappropriate measures of the profit rate, and depend on an implicit assumption that within-industry profit rates are similar in shape.

On each of these points we have shown that empirical testing with UK data confirms that our suspicions are indeed justified.

Our internal critique of Glick used tests based on his own methods: these were shown to produce different estimates of degree of gravitation at different levels of aggregation, as they also do for the whole economy as opposed to manufacturing. Only our relativised modification of Glick's test statistic escapes problems in comparing whole economy gravitation with that within manufacturing, and in comparing UK manufacturing with US manufacturing. However, our version of the statistic shows Glick's preferred measures of the profit rate to provide less evidence of gravitation than the marxist measures, directly contrary to both his theoretical assumptions and his empirical tests.

Our external critique of Glick used tests based on random-sized random samples and the method of *L*-moments to reveal the reason for the results of the preceding work. His profit rate measures show that the overall dispersion of industry profit rates is very small, while the dispersion of industry time average rates is large compared to this overall dispersion; with the marxist measures the industry time averages have a large dispersion, but this is small relative to the overall dispersion. Independently of this, Glick's assumption that convergence of mean rates is a reliable measure of gravitation is shown to be questionable in respect of his preferred profit rate measures (but, ironically, Gillman's marxist ones are less likely to be so), because of the possibility of radically different distributions of profit rates within industries at the two-digit level of aggregation.

Chapter 7 Conclusions

This work makes contributions of both substance and method. The substantial contributions are tests of hypotheses about company profit rate distributions using a large data set, information about the variation of these distributions over time, and the application of Glick's methodology for testing equalisation of profit rates to UK data. Methodological contributions include theoretical and empirical evaluation of alternative definitions of the profit rate, the introduction to economics of recent innovations in moment theory, and the use of random sampling techniques to estimate statistics for profit rate distributions across the total invested capital.

We set out with the primary aim of testing the hypothesis that company profit rates, once weighted by the capital employed, follow a gamma distribution (Farjoun and Machover 1983). The motivation for this was to investigate their claim to have *dissolved* the transformation problem by adapting the concepts and methods of statistical mechanics to Marxist political economy.

Our first task was to argue that company accounts data are intrinsically suitable to the task of estimating the Marxist notion of the rate of profit, and thus against the view that such data cannot measure a meaningful notion of the rate of return. This allowed us to use a large set of company accounts from FAME to compute a number of different definitions of the rate of profit which are discussed in the Marxist empirical literature.

Exploratory data analysis of profit rate distributions at the company level showed that different types of profit rate measure have quite different distributional forms. This also showed that the empirical distributions not only have heavy tails but are susceptible to outliers, and that the range of profit rates observed is inversely related to the size of firm.

L -moments are alternatives to traditional moments which are robust in the presence of outliers yet sensitive to variations in the tails of distributions. We used a size-weighted sampling technique to estimate the L -moment ratios of the distributions of the different profit rates over the capital employed, and hence identify the distributional models which

they follow. This work, and subsequent Zipf plot analysis, showed that while members of the family of exponential distributions (such as the gamma) are strong candidates for modelling the main mass of the data, all profit rate measures have power law tails that dominate the L -moment estimates.

A contemporary alternative to Farjoun and Machover's approach to the transformation problem was that of Glick (1985). We used our data set and Glick's methodology to show that his work is vulnerable to a number of criticisms developed within Glick's own framework, and we used our L -moment and random sampling technique to elucidate the reasons for this. We also used these latter techniques to demonstrate further criticisms advanced from the theoretical perspective developed by Farjoun and Machover.

Despite its promise, the gamma hypothesis cannot be firmly asserted in the absence of further work on the power law tails that we have shown to be prevalent. This further work must either provide ways to distinguish the tails from the main body of the data, or subsume the gamma hypothesis into a more general model. Nor can we dismiss the possibility that some entirely different probabilistic model might provide a unified account of the empirical features we have revealed.

The chapter therefore begins by reviewing our work and noting the methodological innovations adopted; we conclude by pointing out some ways in which the present work might be further developed.

7.1 Contributions

7.1.1 Evaluating profit rate distributions

We set out to test two hypotheses about profit rate distributions: Gibrat's claim (1931) that it should be log normal (across the population of firms), and Farjoun and Machover's claim (1983) that it should be gamma, across the population of money capital invested.

In the case of Farjoun and Machover (but not of Gibrat) their hypothesis concerns a particular definition of the profit rate. Thus our test involved not only attempting to

confirm or reject these particular hypotheses, but also consideration of the distributions of each profit rate definition, at both the firm and capital levels.

Both Gibrat and Farjoun and Machover provide small-scale illustrations to indicate that their respective hypotheses have some foundation in reality. Neither pretends that these amount to systematic tests. This is unsatisfactory in either case, but particularly so in respect of Farjoun and Machover. While better tests of Gibrat's hypothesis might simply seek larger, more representative samples, the whole point of Farjoun and Machover's theory is that it applies to the population of firms as a whole. Because of the large number of firms for which we can calculate the various profit-rate measures (see Chapter Three, Table 3.4) our work constitutes a test on a very much larger scale than any previous attempt.

As our work over Chapters Three, Four and Five showed, these tests are complicated by the presence of heavy tails for all definitions of the profit rate, at both firm and capital levels. These tails are characteristic of those associated with power law distributions such as the Pareto distribution.

Nonetheless, in Chapter Five we found some evidence to identify the particular profit rate favoured by Farjoun and Machover as having a log normal distribution. This distribution is a special case of the generalised gamma distribution, and finding it associated with this profit rate measure – but not with alternative measures – is the case we dubbed 'neutral confirmation' of their hypothesis in Table 5.1. Strong confirmation will depend on further development of methods to deal with the power law tails.

However, the demonstration of power law tails is itself an important discovery, as it aligns the study of profit rate distributions with the growing econophysics literature, discussed in Chapter One, that interprets economic phenomena as the outcome of complex adaptive systems and in which power law tails are found to characterise the distributions of many variables, including firm size as well as income and wealth.

A further contribution is our demonstration in Chapter Three that at the firm level the extent of the tails is clearly related to company size, as measured by the capital concept involved in the relevant profit rate.

7.1.2 Longitudinal variation in profit rate distributions

Variation over the course of the business cycle seems to be a question that occurs naturally to those who take an interest in the distribution of profit rates. The point is raised in Gibrat, and even there it is in the form of a critique of previous hypotheses (page 183). However, the question actually considered is merely the fate of the variance, rather than what might be considered the more sophisticated one of variation in parameters of the distribution. Admittedly, this is not a distinction if one believes that the distribution is log normal. Gibrat concludes that profit rates tend to become more dispersed during the upswing of the cycle (*l'essor augmente l'inégalité des profits, la depression la diminue*).

Farjoun and Machover do consider parameters explicitly. They find, for two selected years, that the (capital-weighted) *average* profit rate *increases* in a recession year compared to a boom year, but the parameters of their fitted gamma distribution change in such a way that the *modal* rate falls. Thus the proportion of total capital earning low rates of return increases markedly. As far as it goes, this contradicts Gibrat's view. Once again, the lack of systematic tests of their work means that no further direct information on this point exists.

However, recent work (Higson *et al.*, 2002, Higson *et al.*, 2004) provides indirect evidence. Using data for the US (1950-1999) and UK (1967-1997) they calculate the skewness and kurtosis, and kernel density estimates (UK only), of growth rates of company sales. They find that kurtosis is positively correlated with the business cycle, and the skewness and standard deviation negatively so. But they do not attempt to model the distributional law, and hence are unconcerned with parameters, as opposed to moments, of the distribution.

The *L*-moment methods we used to identify possible distributional models depend on examination of patterns of annual variation in *L*-skewness and kurtosis, from which

estimated parameters can be easily derived once a particular model is selected. However our work in Chapter Five demonstrated that the variation in the L -moment ratios is dominated by the power law tails of the empirical distributions, so much so that for most profit rate measures the possible models appeared to be outside the range of distributions for which Hosking and Wallis (1997) provide methods for estimating the parameters.

Once again this unanticipated outcome appears to be of potential significance. Given the inverse relation between size of firm and propensity to extreme values of the profit rate, the conclusion is what everyday experience suggests: that the fortunes of small firms are particularly sensitive to the business cycle.

7.1.3 Competition as gravitation

Our interest in profit rate distributions stemmed from a view that profit rate equalisation, supposedly an indispensable axiom in Marx's transformation problem, was not only questionable on independent methodological grounds but also inconsistent with Marx's own outlook.

Glick's interpretation of classical political economy's notion of competition is that although industry profit rates vary through time, the average deviation from the overall profit rate will be small (by implication, small in comparison to the dispersion within industries). Although deviation is a persistent fact, the actual values of industry rates of profit will 'orbit' (in Glick's phrase) a common centre of gravity. He tests this 'gravitation' of profit rates using a variety of techniques, but the aspect of his work that most closely parallels our own is a set of tests using data derived from company accounts, and it is natural to test it by applying Glick's methods to our data.

In Chapter Six we have done much more than simply replicate Glick's work in a different setting. Our methodological perspective led us to an extensive critique of Glick on both internal and external criteria, which in turn suggested extensions and modifications to Glick's procedures, made possible by a data set which is both more extensive and more detailed than that used by Glick.

The internal critique of Glick consisted of five points (the first made in Chapter Two, the rest in Chapter Six): that his preferred definition of the rate of profit is not appropriate to the notion of the competitive process that is to be tested; that he narrowly confines his tests to the manufacturing sector rather than to the economy as a whole; that his empirical work is at too high a level of aggregation to accord with his notion of ‘industry’; that his choice of measure of gravitation can be questioned; and that the full implications of his hypothesis are not tested. The external critique (Chapter Six) was that no attention is paid to the shape of profit rate distributions, only to their dispersion.

On profit rate definitions, Glick prefers measures which include financial assets, yet the passage from Ricardo which he uses to illustrate the classical conception of the competitive process appears to suggest that a measure of the value of physical assets alone would be more appropriate. On the level of aggregation, Glick himself suggests that two-digit industry data is likely to be a poor choice for testing a theory whose notion of industry is closer to that of ‘line of business’.

Glick’s test of gravitation is the absolute value of a measure (V_1) of the total dispersion of each industry’s long-run average profit rate from the general rate of profit, but this takes no account of how large this Figure is in relation to the total variation in industry profit rates over time (V). We thus proposed a relativised version of this, $\hat{V} = V_1/V$.

The final point of the internal critique was that although Glick’s account implies that *inter*-industry profit rate dispersions should be much smaller than *intra*-industry ones, this corollary is not tested.

Our external critique was that no sustained attention is paid to the shapes of distributions, in the sense that moments higher than the second are scarcely discussed. Coupling this point to the final part of the internal critique, we argued that focussing solely on the mean of industry profit rates inevitably risked ambiguous or misleading results if skewed within-industry profit rate distributions were prevalent.

To test the internal critique we applied Glick's methods to our FAME data set and found that tests of gravitation in the manufacturing sector alone gave puzzlingly different results to those for the economy as a whole, although the expected relationship between degree of aggregation and degree of gravitation was found. The puzzles in the manufacturing/whole economy comparison are resolved if our relativised version of the test statistic is employed, at the expense of introducing a perverse relation between aggregation and gravitation.

The foregoing results all relate to the profit rate definitions preferred by Glick but criticised by us. But adopting the definitions which we identified in Chapter Two as being the appropriate ones for testing Farjoun and Machover's very different perspective, together with our relativised test statistic, not only restores the expected relationship between aggregation and gravitation, but provides superior evidence of gravitation.

Further, an investigation using the methods developed to identify profit rate distributions readily revealed why Glick's qualitative conclusions proved vulnerable: the data proved to have exactly the feature which our external critique pointed out to be capable of vitiating Glick's methods, namely widespread skewness of within-industry distributions, which moreover could be of either sign. The implication is that considering the dispersion of industry means, rather than some other measure of central tendency such as the mode, may be a very unreliable method of assessing gravitation.

7.1.4 *L*-moments in economics

Several problems arose very quickly in our exploratory data analysis. Recall that we used our data to test hypotheses about two different sample spaces: firms, and the capital employed. The following discussion assumes that we are working in the firm space in the first instance.

First was the existence of numerous extreme values, too closely spaced to be dismissed as outliers (itself a problematic approach), especially as some of the candidate distributions are known to be outlier-prone.

Second, for some profit-rate measures, was the existence of negative values (including extreme negative values). But many of the candidate distributions are bounded, at zero in their standard forms. This is not problematic as such, either economically or mathematically. From an economic viewpoint, extreme values can be shown to be associated with very small firms, which are unlikely to be genuinely capitalist enterprises. This leaves the lower bound as a measure of the minimum rate of profit (or, more intuitively, maximum rate of loss) which the capital market will permit a firm to sustain. Estimating this is clearly an interesting project. Mathematically, one can simply introduce a location parameter to shift the distribution an appropriate distance along the x -axis; however, including this further parameter increases the technical problems of estimation.

Third, judging whether distributional models are supported by the data is normally done by goodness-of-fit tests such as the Kolmogorov-Smirnov test (Kolmogorov, 1992). But these present the difficulty that the larger the data set, the more likely it is that any test will indicate a poor fit. One could of course sample the data – but lacking a clear *a priori* basis for choosing the sample size one runs the risk of choosing the size so as ensure a fit. This problem might be solved by bootstrapping procedures (Efron and Tibshirani, 1993), but these do not avoid the second problem above, that of estimating the location parameter. In any case, we were attempting to estimate the population density directly.

All three issues can in principle be addressed by selecting and estimating distributional models using L -moments (Hosking and Wallis, 1997), an alternative to the usual Pearsonian product-moments. They avoid excessive sensitivity to extreme values and thus allow more accurate discrimination between distributions that differ mainly in the characteristics of their extreme tails.

Further attractive properties of L -moments are that all moments are guaranteed to exist if the mean does, that their estimates have low bias, and that no two distributions share the same set of L -moments. Further, L -moment ratios, unlike Pearsonian moment ratios, are unbounded, and the values of sample statistics are not constrained by sample size from taking on any value attainable by population statistics. These facets make the method of L -

moments reliable in identifying distributions and estimating their parameters, in contrast to the method of traditional moments.

For our purposes they have the further advantage of providing location estimates for bounded distributions simply as polynomial functions of the sample L -moments.

These properties proved invaluable in the face of the unexpected results in Chapter Five, in which we used a quasi-bootstrapping procedure to estimate the L -moments to obtain size-weighted samples from which to estimate distributions in the capital space. These were that power law tails of the distributions of all profit rate measures dominate the variation in estimated L -skewness and L -kurtosis, and in some cases produce L -kurtosis well in excess of that associated with any of the distributions for which the loci in L -skewness/ L -kurtosis space are given by Hosking and Wallis (1997). Only one profit rate definition, Gillman 4, lay within the band of loci available, and the axis of this measure's L -moment ratios was sufficiently weakly aligned with the relevant loci to suggest caution in asserting any particular model.

Because we could be confident that the observed L -moment ratios depicted real features of the data, provided that they had been correctly estimated, we were prompted to consider the L -moments of the individual samples in our random-sized random sampling procedure, and this pointed the way to a re-examination of the tails of the capital-weighted distributions and discovery of the power law tails.

7.1.5 Choice of profit rate definition

The empirical Marxian literature on the profit rate is characterised by controversy over the proper definition of this concept. As our review in Chapter Two showed, the views of different authors vary partly because of different interpretations of Marx, and partly because they address themselves to different problems.

Thus writers primarily interested in the long-run movement of the general rate of profit have tended to advocate measures that take account of expenditure on (and in some cases,

of investment in) unproductive activities. Glick's work on gravitation assumes that the profit rate to use is that achieved by industrial capital not only after unproductive expenditure, but after all distribution to other claimants. In contrast, we argued that Farjoun and Machover's gamma hypothesis related to the definitions Gillman (1957) dubbed 'marxian', in distinction to the 'capitalist' measures involving unproductive expenditure.

Our data set allowed us to compute no less than 17 different measures identified in this literature, together with four standard accounting ratios, and to test all these in the context of both Farjoun and Machover and Glick, two important but very different recent responses to the Marxian transformation problem.

In Chapter Five we found evidence that one of Gillman's Marxian measures (Gillman 4, which may equally well be dubbed the Farjoun and Machover measure) does indeed have a gamma distribution, once that is taken to include the four-parameter generalisation of this distribution rather than the two-parameter shape and scale version discussed by Farjoun and Machover themselves. Since the particular variant of the four-parameter gamma found is the log normal distribution, this is also evidence in favour of Gibrat's hypothesis about profit rates, provided that it is interpreted as applicable to the capital-level distribution and not solely to the firm-level distribution. This evidence might well be strengthened if the power law tails were to be separately accounted for.

Our empirical testing thus showed that it is the Marxian measures and not their rivals which most plausibly instantiate the distributional hypotheses under test. In contrast, the 'capitalist' measures of Gillman and Glick cannot easily be thought of as having gamma distributions, or any other in the band of Hosking and Wallis loci, at any rate pending some separate treatment of their power law tails.

In Chapter Six we went on to show that Gillman's 'marxian' measures can be argued to be better measures, in the context of Glick's work on gravitation, than Glick's own candidates. Replicating and extending Glick's empirical work showed that narrow (or

‘capitalist’) measures do not demonstrate the validity of his conception of the competitive process, notwithstanding his own results: in fact, they lead to contradictory results which the Marxian measures avoid.

Before doing so any of this work, however, we had to rebut the claim that accounting data is intrinsically incapable of providing information about the rate of return. We did so in two ways. First, we appealed to Kaldor’s argument that the subjective approach to the concepts of income and capital was at least as incoherent as the objective, and to Simons’ argument that the logic of income was one of *gain*, not a flow of enjoyments. Second, we adopted Bryer’s argument that traditional accounting data is precisely what social capital (the collective investor constituted by the capitalist class) uses to monitor business managers’ stewardship of their wealth – in particular, Simons’ ‘gain’: the rate of increase in wealth, or profit rate.

However, the most interesting qualitative result may be the clear evidence that the many different definitions we test fall into a relatively small number of groups, viewed in the light of the empirical densities presented in Chapters Three and Four.

The interesting question is *why* do different profit rate definitions result in different distributions? The technical answer is that profit rate distributions are convolutions of the distribution of net income with that of capital, and as we have seen, it is the latter distribution that appears to dominate the results.

However, from a theoretical point of view we suggest the following for consideration: the Marxian measures we have tested are, arguably, measures of the rate of surplus value created in production; the capitalist ones are measures of the rate of surplus appropriated as a result of the *total* process of circulation, including the distribution of the surplus in accordance with the claims implied by firms’ financial structures.

But converting the non-equalisation of the former – the rate of surplus from production – to the equalisation of the latter – the rate of profit on capital as perceived by capitalists and their agents – is what the transformation problem is supposed to be about.

7.3 Directions for future research

The demonstration of extensive power law tails for a wide variety of profit rate definitions means that strategies for investigation in this area must be rethought. There would seem to be two broad alternatives.

First, one might amend the existing hypotheses of Farjoun and Machover and Gibrat so as to refer only to the main mass of the data, and seek some theoretical model that explained why different distributional models should be expected for the main mass on the one hand and the power law tails on the other.

Secondly, one might look for an alternative theoretical model which gave rise to a single distributional model for the whole of the data (the order of actual work might of course be different; one might search for a single descriptive law and then attempt to provide a theoretical explanation for it).

In the first case one might seek to develop methods for identifying the boundary between the power law tail and the main mass, with a view to identifying and estimating the distributions governing each part. Such an approach would, preferably, be accompanied by development of a theoretical explanation in which such a mixture of distributions would be an expected feature of actual data. Hypotheses such as those of Gibrat and Farjoun and Machover would then necessarily be recast in more restrictive form. A starting point for this line of attack would be the literature dealing with the similar features found in the distributions of company size, and of individual wealth and income.

In the second case one would look for an alternative model that included the tails of the data as an intrinsic feature. Again, it would be desirable for the model to be justified on theoretical grounds, rather than adopted purely for descriptive purposes.

One possible candidate for such a model is the stable family of distributions. This has the two-fold attraction of promising a flexible descriptive model that can accommodate very heavy tails over the whole real line, while also accommodating variants with bounded support, at the same time as potentially linking the theory of company profit rates from production with that of market returns to financial assets.

Finally, there are some obvious questions of detail which the techniques used in this study might illuminate. It would, for example, be natural to extend the tentative study of inter-industry differences in distribution which we made in the final section of Chapter Six. Equally, a more detailed study of how profit rate distributions vary among size classes of firm would not only be a potentially useful contribution to the literature on small and medium enterprises, but would play an important part if the mixture-of-distributions approach were to be pursued.

Appendix

This appendix deals with some technical features of gamma distributions. The two-parameter gamma is the distribution proposed by Farjoun and Machover (1983) for the profit rate, as weighted by company size (see section 5.1.1). Not only is this the core hypothesis tested in the current work, but the assumption that this distribution obtains is also the key to a number of other features of their work, such as their argument that the only possible uniform rate of surplus value is unity.

We start with the two-parameter case envisaged by Farjoun and Machover, note the existence of an alternative parameterisation, and consider its generalisation with three and four parameters; we then demonstrate the relation of the latter to other types of two-parameter (shape and scale) distributions, among which is the log normal distribution proposed by Gibrat (1931) for the profit rate distribution. We then review some issues in estimation of parameters of gamma distributions and other tests involving these distributions.

Probability density function

Two-parameter (scale and shape)

In the parameterisation used by Farjoun and Machover the probability density function (pdf) is given by

$$f(X; \lambda, \alpha) = \frac{\lambda^\alpha}{\Gamma(\alpha)} X^{\alpha-1} e^{-\lambda X}$$

where $\Gamma(\alpha) = \int_0^\infty e^{-u} u^{\alpha-1} du$ is the gamma function, $\alpha > 0$, $\lambda > 0$ and $0 < X < \infty$; α is a shape parameter, and λ is generally referred to as a ‘rate’ parameter (Salem and Mount, 1974).

By substituting $\alpha = \frac{1}{\theta}$ in the above one obtains

$$f(X; \theta, \alpha) = \frac{X^{\alpha-1} e^{-X/\theta}}{\Gamma(\alpha)} \text{ where } \theta \text{ is a scale parameter.}$$

Figure A.1 illustrates the latter parameterisation with shape parameters 0.5, 1, 1.5, 2 and 3 and scale = 1. When $\alpha < 1$ the density has a pole at $X = 0$; when $\alpha = 1$, $f(X=0) = 1$, and for $\alpha > 1$ the density is zero.

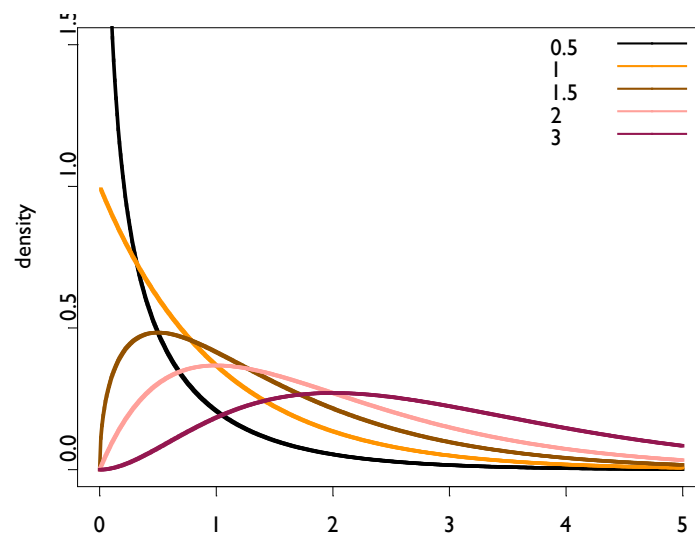


Figure A.1: gamma distributions; shape parameters as indicated in legend

Three-parameter (location, scale and shape)

The two-parameter gamma is only defined for non-negative X ; but the above is easily extended to cover the case of variables capable of taking negative values (or indeed with strictly positive minimum values) by adding a threshold parameter μ :

$$f(X; \mu, \theta, \alpha) = \frac{(X - \mu)^{\alpha-1} e^{-(X - \mu)/\theta}}{\Gamma(\alpha)} \text{ or } f(X; \mu, \theta, \alpha) = \frac{(X - \mu)^{\alpha-1} e^{-(X - \mu)/\theta}}{\Gamma(\alpha)}.$$

Hosking and Wallis point out (1997:200–201) that re-parameterising with the conventional moments of the distribution as location-scale-shape parameters μ, σ, α , where $\sigma^2 = 4/\alpha^2$, $\mu = \frac{1}{2} \left(1 + \frac{1}{\alpha} \right)$ and $\alpha = 2/\sigma^2$, allows for both positively and negatively skewed versions, depending on the sign of σ .

Four-parameter (location, scale, shape and power/shape)

If the random variable Z has the standard gamma distribution (that is with scale parameter $= 1$ and $= 0$) and $Z = \{(X - \mu)/\sigma\}$ then the four-parameter, or generalised, gamma distribution has pdf

$$f(X; \mu, \sigma, \alpha, \beta) = \frac{(X - \mu)^{\alpha-1} e^{-(X - \mu)/\sigma}}{\sigma^\alpha \Gamma(\alpha)}$$

where β is a second shape parameter (Johnson *et al.*, 1994: 388).

The generalised gamma and other distributions

McDonald (1984: 648) shows that the generalized gamma (GG) is itself a special case of both of the generalized betas of the first and second kind (GB1, GB2), where GG, GB1 and GB2 are defined by (1), (2) and (3), respectively (omitting location parameters):

$$\begin{aligned} (1) \quad f(y; a, p) &= \frac{ay^{ap-1} e^{-(y/b)^a}}{b^{ap} \Gamma(p)}, & 0 < y, \\ (2) \quad g(y; a, b, p, q) &= \frac{ay^{ap-1} (1 - (y/b)^a)^{q-1}}{b^{ap} B(p, q)}, & 0 < y < b, \\ (3) \quad h(y; a, b, p, q) &= \frac{ay^{ap-1}}{b^{ap} B(p, q) (1 - (y/b)^a)^{q-1}}, & 0 < y, \\ &= 0 & \text{otherwise} \end{aligned}$$

where the beta function $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ and $\Gamma(\cdot)$ is the gamma function.

In (2), letting $q \rightarrow \infty$ and setting $b = (p+q)^{1/a}$ produces (1).

In (3), letting $q \rightarrow \infty$ and setting $b = q^{1/a}$ produces (1).

Figure A.2 reproduces Figure 1 in McDonald (1984). As well as showing the relationships between GB1, GB2 and GG, it shows how GG has as special or limiting cases

not only the two-parameter gamma (GA) but also the Weibull (W) and the log normal (LN) distributions.⁷⁹

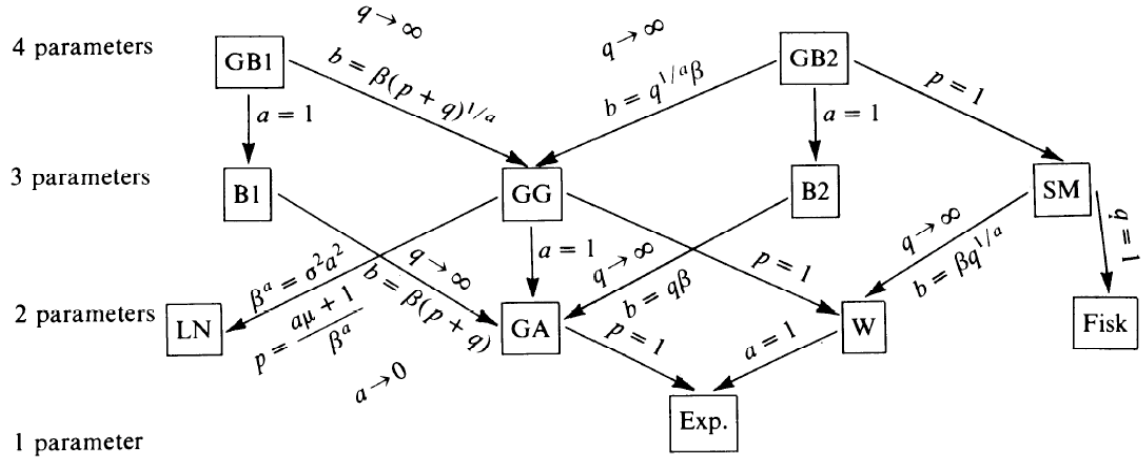


Figure A.2: relationship of generalised gamma to other distributions (from McDonald 1984)

Other special or limiting cases of GG include the χ^2 and Pareto distributions (Evans *et al.*, 1993: 80–81), the Rayleigh distribution, and the half-, circular and spherical normal (Stacy and Mihram, 1965: 351); the latter lead directly to the Maxwell molecular speed and velocity (and hence energy) distributions, to which Farjoun and Machover (1983) make heuristic appeal.

Taguchi (1980) re-parameterises the generalised gamma as:

$$f(X; \gamma, h) = \frac{X^{\gamma-1}}{(h^{\gamma-1} + 1)} \exp(-X^h)$$

where γ is skewness ($=1/c$, where c is the shape parameter of the usual gamma distribution) and h is ‘sharpness’ ($=cd$, where c, d are shape parameters in the usual version of the two-shape-parameter distribution).

⁷⁹ The other distributions shown are the beta distributions of the first and second kind (B1 and B2), the Singh-Maddala (SM), Fisk, and exponential (Exp.).

Other properties

Characterisation

The gamma distribution has the following unique property: if R and Z are independent non-degenerate random variables, then the random variables $R+Z$ and R/Z are also independent if and only if R and Z have gamma distributions with a common scale parameter (Marsaglia, 1974, 1989). This is an extension of Lukacs (1955), which is cited by Farjoun and Machover in their proof that R/Z must necessarily be equal to one, if it is indeed single-valued (1983:204).

Being unique to the gamma, this property provides a basis for testing for whether data is gamma-distributed; see Shapiro (2001) for a discussion of methods.

Inequality measures

In the case of the two-parameter gamma, the shape and scale parameters and are easily linked to three standard measures of inequality, the Lorenz, Gini and Theil indices (see Salem and Mount, 1974: 1116–1117).

Lorenz concentration ratio $L = 2B_{0.5}(\alpha, \alpha+1) - 1$ where $B_{0.5}(\alpha, \alpha+1)$ is the incomplete beta function.

Gini co-efficient of mean difference $G = 2L / \pi$.

Theil index $I = \frac{1}{\alpha} + \frac{1}{\alpha} \log \alpha$, where $\psi(\alpha) = \psi'(\alpha) / \Gamma(\alpha) = d \log \Gamma(\alpha) / d\alpha$ is the digamma function.

Reproductive property

If X_i are independent gamma variables with common and parameters but not necessarily different α_i , then $\sum_{i=1}^n X_i$ has a gamma distribution with parameters α , and

$$= \sum_{i=1}^n \alpha_i \quad (\text{Johnson } et \text{ al., 1994: 340}).$$

Estimation

Estimation of the generalised gamma with four parameters unknown is not straightforward. As might be expected from Taguchi's derivation of a 'sharpness' parameter as the product of the two usual shape parameters, a fundamental problem is that maximum likelihood estimation (MLE) of these leads to highly negatively correlated results (Johnson *et al.*, 1994: 393).

One might consider estimating the parameters of the Taguchi version and recovering the conventional first shape parameter as the inverse of Taguchi's skewness parameter, and the second conventional shape parameter as $h/$, where h is Taguchi's sharpness parameter. However, Taguchi's estimation procedure relies on the relationship

$$\frac{\frac{2}{l}}{\frac{2}{l}} = \frac{\frac{1}{h}}{\frac{1}{h}}, \quad = \frac{1}{l} \quad \frac{1}{h}$$

where l and l^2 are the logarithmic mean and logarithmic variance respectively and $\psi(x)$ is the di-gamma function. It is unclear whether the resulting estimators would have desirable properties.

In any case it is debatable whether attempting to estimate the generalised gamma distribution is worthwhile, even for purely descriptive purposes. Given that particular cases of it are easily estimated by the method of L -moments discussed in Chapter Four it would seem more useful to select the appropriate location-scale-shape distribution using L -moment ratio diagrams and then estimate those. If further theoretical development justifies use of the four-parameter gamma (particularly if it suggests definite economic meanings for the parameters) then the matter can be revisited.

Given the simplicity and reliability offered by the method of L -moments, MLE should be avoided in the first instance, even in the case of the location-scale-shape gamma. It is the estimation of the threshold parameter that creates difficulties, as MLE homes in on the

minimum observation in the sample for its estimate, but this choice leads to inconsistent ML estimates of the other parameters (Cheng and Iles, 1987); see also Cheng and Traylor (1995).

Hirose (1995) discusses a ‘predicto-corrector’ MLE method in this case, but it is complex and can result in multiple local estimates of the shape parameter. Other MLE problems are discussed in McCullagh and Nelder (1989: 295–296); other methods are discussed in Johnson *et al* (1993).

Bibliography

- Amoroso, L. (1925) 'Ricerche intorno alla curva dei redditi', *Annali di Matematica Pura ed Applicata*, **Ser. 4-21**, pp. 123-159.
- Anderson, T.W. and Darling, D.A. (1952) 'Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes', *Annals of Mathematical Statistics*, **23:2**, p. 193-212.
- Ashley, W.J. (1910) 'The statistical measurement of profit', *The Economic Journal*, **20:80**, pp. 530-550.
- Axtell, Robert (2005) 'The complexity of change', *The Economic Journal*, **115:June**, pp. F193-F210.
- Barnett, V. and Lewis, T. (1994) *Outliers in statistical data*, John Wiley and Sons.
- Braun, Dieter (2006) 'Nonequilibrium thermodynamics of wealth condensation', *Physica A: Statistical and Theoretical Physics*, **369:2**, pp. 714-722.
- Bryer, R.A. (1993) 'The late nineteenth century revolution in financial reporting: accounting for the rise of investor or managerial capitalism?' *Accounting, Organizations and Society*, **18:7/8**, pp. 649-690.
- Bryer, R.A. (1994) 'Why Marx's labour theory is superior to the marginalist theory of value: the case from modern financial reporting', *Critical Perspectives on Accounting*, **5:4**, pp. 313-340.
- Bryer, R.A. (1999) 'Marx and accounting', *Critical Perspectives on Accounting*, **10**, p. 683-709.
- Bryer, R.A. (2000) 'The history of accounting and the transition to capitalism in England. Part one: theory', *Accounting, Organizations and Society*, **25:25**, pp. 131-162.
- Bureau of Economics (1981) 'Statistical report: annual line of business report 1974 (interim)', Federal Trade Commission, Washington DC.
- Champernowne, D.G. (1985) 'Review of *Laws of Chaos*', *British Journal for the Philosophy of Science*, **36:3**, pp. 348-352.
- Cheng, R.C.H. and Iles, T.C. (1987) 'Corrected maximum likelihood in non-regular problems', *Journal of the Royal Statistical Society, series B*, **49:1**, pp. 95-101.
- Cheng, R.C.H. and Traylor, L. (1995) 'Non-regular maximum likelihood problems', *Journal of the Royal Statistical Society, series B*, **57:1**, p. 3-44.
- Christodouloupoulos, George (1996) 'International competition and industrial rates of return: an empirical approach 1970-1990', *IWGVIT*, Boston.
- Cockshott, Paul and Cottrell, Allin (1994) 'Value price transformation as a real process', *Karl Marx's third volume of 'Capital': 1894-1994*, University of Bergamo.
- Cockshott, Paul and Cottrell, Allin (1998) 'Does Marx need to transform?' in *Marxian economics: a reappraisal*, Vol. 2 (Ed, Bellofiore, Riccardo) Macmillan, London, pp. 70-85.
- Cockshott, Paul and Cottrell, Allin (2003) 'Robust correlations between prices and labour values: a comment', *Cambridge Journal of Economics*, **29**, pp. 309-316.
- Cockshott, Paul, Cottrell, Allin and Michaelson, Greg (1995) 'Testing Marx: some new results from UK data', *Capital & Class*:55, pp. 103-129.
- Creedy, John and Martin, Vance L. (1994) 'A model of the distribution of prices', *Oxford Bulletin of Mathematics and Statistics*, **56**, pp. 67-76.
- Creedy, John, Lye, Jenny N. and Martin, Vance L. (1994) 'A labour market equilibrium model of the personal distribution of earnings', University of Melbourne, Department of Economics, Melbourne.
- DeLamarter, Richard Thomas (1987) *Big Blue: IBM's use and abuse of power*, Macmillan, London.
- Dermietzel (1906) *Statistische Untersuchungen uber die Kapitalrente der grosseren Aktiengesellschaften von 1876-1902*, Gottingen.

- Doodson, Arthur T. (1917) 'Relation of the mode, median and mean in frequency curves', *Biometrika*, **11**:4, pp. 425-429.
- Duménil, Gérard and Lévy, Dominique (1987a) 'The dynamics of competition: a restoration of the classical analysis', *Cambridge Journal of Economics*, **11**:2, pp. 133-164.
- Duménil, Gérard and Lévy, Dominique (1987b) 'The trend of the economic rate of return in the [sic] U.S. manufacturing since the depression', CEPREMAP, Paris.
- Duménil, Gérard and Lévy, Dominique (1990) 'Post depression trends in the economic rate of return for U.S. manufacturing', *The Review of Economics and Statistics*, **72**:3, pp. 406-413.
- Duménil, Gérard and Lévy, Dominique (1994) 'Stylized facts about technical progress since the Civil War: a vintage model', *Structural Change and Economic Dynamics*, **5**:1, pp. 1-23.
- Duménil, Gérard and Lévy, Dominique (1999a) 'The acceleration and slowdown of technical progress in the US since the Civil War: the transition between two paradigms', CEPREMAP.
- Duménil, Gérard and Lévy, Dominique (1999b) 'Profit rates: gravitation and trends', CEPREMAP.
- Duménil, Gérard and Lévy, Dominique (1999c) 'The profit rate: where and how much did it fall? Did it recover? (USA 1948-1997)', CEPREMAP.
- Dunne, Paul (1991) 'An introduction to quantitative Marxism', in *Quantitative Marxism*, (Ed, Dunne, Paul) Polity Press in association with Basil Blackwell, Cambridge.
- Edgeworth, F.Y. (1883) 'The method of least squares', *Philosophical Magazine*, **16**, p. 78-84.
- Edwards, Jeremy, Kay, Jon and Mayer, Colin (1987) *The economic analysis of accounting profitability*, Clarendon Press, Oxford.
- Efron, Bradley and Tibshirani, Robert J. (1993) *An introduction to the bootstrap*, Chapman and Hall, London.
- Engels, Friedrich (1978 [1884]) 'Preface', in *Capital*, Vol. 2 (Ed, Mandel, Ernest) Penguin, Harmondsworth, p. 83-102.
- Engels, Friedrich (1981 [1895]) 'Preface', in *Capital*, Vol. 3 (Ed, Mandel, Ernest) Penguin, Harmondsworth, pp. 91-111.
- Evans, Merran, Hastings, Nicholas and Peacock, Brian (1993) *Statistical distributions*, John Wiley & Sons, New York.
- Farjoun, Emmanuel (1984) 'The production of commodities by means of what?' in *Ricardo, Marx, Sraffa*, (Eds, Mandel, Ernest and Freeman, Alan) Verso, London, pp. 11-41.
- Farjoun, Emmanuel and Machover, Moshé (1983) *Laws of chaos: a probabilistic approach to political economy*, Verso, London.
- Farjoun, Emmanuel and Machover, Moshé (1985) 'Probability, economics and the labour theory of value', *New Left Review*, **152**, July/August, p. 264.
- Fisher, Franklin M. and McGowan, John J. (1983) 'On the misuse of accounting rates of return to infer monopoly profits', *The American Economic Review*, **73**:1, pp. 82-97.
- Fisher, Franklin M., McGowan, John J. and Greenwood, Joen E. (1983) *Folded, spindled, and mutilated: economic analysis and U.S. v. IBM*, The MIT Press, Cambridge MA.
- Fisher, Irving (1969) 'Income and capital', in *Readings in the concept and measurement of income*, (Eds, Parker, R.H. and Harcourt, Geoffrey C.) Cambridge University Press, Cambridge, pp. 33-53.
- Foley, Duncan (1989) 'Notes on ideology and methodology', <http://cepa.newschool.edu/~foleyd/ideo.pdf>; accessed on 31 October, 2006.
- Freeman, Alan (1995) 'Marx without equilibrium', *Capital and Class*, **56**, pp. 49-90.
- Freeman, Alan (1998) 'The transformation of prices into values', in *Marxian economics: a reappraisal*, Vol. 2 (Ed, Bellofiore, Riccardo) Macmillan, London, pp. 270-275.

- Freeman, Alan (2006) 'How economies fluctuate: the ambivalent metaphor of 'centre of gravity' in recent economic writings', *Association for Heterodox Economics annual conference*, London.
- Freeman, Alan and Carchedi, Guiglielmo (1996) *Marx and non-equilibrium economics*, Edward Elgar.
- Freeman, Alan, Kliman, Andrew and Wells, Julian (Eds.) (2004) *The new value controversy and the foundations of economics*, Edward Elgar, Cheltenham.
- Fritsche, Stephen R. and Dugan, Michael T. (1997) 'A simulation-based investigation of errors in accounting-based surrogates for internal rate of return', *Journal of Business Finance and Accounting*, 24:6, pp. 781-802.
- Fujiwara, Yoshi, Aoyama, Hideaki, Di Guilmi, Corrado, Souma, Wataru and Gallegati, Mauro (2004) 'Gibrat and Pareto-Zipf revisited with European firms', *Physica A: Statistical Mechanics and its Applications*, 344:1-2, p. 112-116.
- Fujiwara, Yoshi, Souma, Wataru, Aoyama, Hideaki, Kaizoji, Taisei and Aoki, Masanao (2003) 'Growth and fluctuations of personal income', *Physica A: Statistical Mechanics and its Applications*, 321:3-4, pp. 598-604.
- Gale, B. and Branch, B. (1982) 'Concentration versus market share: which determines performance and why does it matter?' *The Antitrust Bulletin*:Spring, pp. 83-105.
- Gibrat, Robert Pierre Louis (1930) 'Une loi des répartitions économiques: l'effet proportionnel', *Bulletin Statistique générale Française*, 19, p. 469.
- Gibrat, Robert Pierre Louis (1931) *Les inégalités économiques*, Sirey, Paris.
- Gillman, Joseph M. (1956) *The falling rate of profit: Marx's law and its significance to twentieth-century capitalism*, Dennis Dobson, London.
- Glaisher, J.W.L. (1872-3) 'On the rejection of discordant observations', *Monthly Notices of the Royal Astronomical Society*, 33:251.
- Glick, Mark (1985) 'Competition vs. monopoly; profit rate dispersion in U.S. manufacturing industries', PhD dissertation, New School for Social Research, New York NY.
- Glick, Mark and Erbar, H. (1988) 'Profit rate equalization in the U.S. and Europe: an econometric investigation', *European Journal of Political Economy*, 4:1, pp. 179-201.
- Glick, Mark and Erbar, H. (1990) 'Long-run equilibrium in the empirical study of monopoly and competition', *Economic Enquiry*, 28:1, pp. 151-162.
- Glick, Mark and Campbell, D.A. (1995) 'Comment: Classical competition and the compatibility of market power and uniform rates of profit', *Review of Radical Political Economy*, 27:2, pp. 124-153.
- Green, R.F. (1976) 'Outlier-prone and outlier-resistant distributions', *Journal of the American Statistical Association*, 71, p. 502-505.
- Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. (1979) 'Probability-weighted moments: definition and relation to parameters of several distributions expressible in inverse form', *Water Resources Research*, 15, p. 1049-1054.
- Griffin, P.A. (1982) 'Usefulness to investors and creditors of information provided by financial reporting: a review of empirical accounting research', Financial Accounting Standards Board, Stamford, Connecticut.
- Harcourt, G.C. (1969) 'The accountant in a golden age', in *Readings in the concept and measurement of income*, (Eds, Parker, R.H. and Harcourt, Geoffrey C.) Cambridge University Press, Cambridge, pp. 127-131.
- Härdle, W. (1991) *Smoothing techniques: with implementation in S*, Springer, New York.
- Higson, C., Holly, Sean and Kattuman, Paul (2002) 'The cross-sectional dynamics of the US business cycle: 1950-1999', *Journal of Economic Dynamics and Control*, 26:9-10, pp. 1539-1555.

- Higson, C., Holly, S., Kattuman, P. and Platis, S. (2004) 'The business cycle, macroeconomic shocks and the cross-section: the growth of UK quoted companies', *Economica*, **71**:281, p. 299–318.
- Hirose, H. (1995) 'Maximum likelihood parameter estimation in the three-parameter gamma distribution', *Computational Statistics and Data Analysis*, **20**, pp. 343–354.
- Hosking, J.R.M. (1990) 'L-moments: analysis and estimation of distributions using linear combinations of order statistics', *Journal of the Royal Statistical Society Series B*, **52**:1, pp. 105–124.
- Hosking, J.R.M. and Wallis, J.R. (1997) *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press, Cambridge.
- International Institute for Social History (2006) 'Karl Marx/Friedrich Engels Papers', <http://www.eesg.nl/archives/en/files/m/10760604full.php>; accessed on 26 August 2006.
- Jalali-Naini, Ahmad R. (1986) 'Laws of Chaos: a probabilistic approach to political economy, by Emmanuel Farjoun and Moshe Machover', *The Journal of Peasant Studies*, **13**:4, pp. 275–278.
- Jaynes, E.T. (1991) 'How should we use entropy in economics?' unpublished MS.
- Johnson, Norman L., Kotz, Samuel and Balakrishnan, N. (1994) *Continuous univariate distributions*, Vol. 1, John Wiley & Sons.
- Jones, M.C. and Faddy, M.J. (2003) 'A skew extension of the t distribution, with applications', *Journal of the Royal Statistical Society Series B*, **65**:1, pp. 179–194.
- Julius, A.J. (2005) 'Stable great ratios and stationary dispersion of profit rates in unbalanced growth', New School for Social Research, New York.
- Kaldor, N. (1969) 'The concept of income in economic theory', in *Readings in the concept and measurement of income*, (Eds, Parker, R.H. and Harcourt, Geoffrey C.) Cambridge University Press, Cambridge, pp. 123–124.
- Kapteyn, J.C. (1903) *Skew frequency curves in biology and statistics*, Nordhoff, Groningen.
- Kliman, Andrew J. (2002) 'The law of value and laws of statistics: sectoral values and prices in the US economy, 1977–97', *Cambridge Journal of Economics*, **26**, pp. 299–311.
- Kliman, Andrew J. (2003) 'Reply to Cockshott and Cottrell', *Cambridge Journal of Economics*, **29**, pp. 317–323.
- Kolmogorov, Andrei Nikolaevich (1992) 'On the empirical determination of a distribution law', in *Selected works of A.N. Kolmogorov*, Vol. Two (Eds, Prokhorov, Yu. V. and Shiryaev, A.N.) Kluwer, Dordrecht, p. 139–146.
- Langston, Robert (1984) 'A new approach to the relation between prices and values', in *Ricardo, Marx, Sraffa*, (Eds, Mandel, Ernest and Freeman, Alan) Verso, London, pp. 1–10.
- Lévy, Dominique (1984) 'A method of dispersion measurement', CEPREMAP, Paris.
- Lianos, Theodore P. and Droucopoulos, Vassilis (1993) 'Convergence and hierarchy of industrial profit rates: the case of Greek manufacturing', *Review of Radical Political Economics*, **25**:2, pp. 67–80.
- Lienhard, John H. and Meyer, Paul L. (1967) 'A physical basis for the generalized gamma distribution', *Quarterly Journal of Applied Mathematics*, **5**:3, pp. 330–334.
- Lukacs, Eugene (1955) 'A characterization of the gamma distribution', *Annals of Mathematical Statistics*, **26**, pp. 319–324.
- Mage, Shane Henry (1963) 'The 'law of the falling tendency of the rate of profit': its place in the Marxian theoretical system and relevance to the U.S. economy', PhD, Columbia, New York NY.
- Maldonado-Filho, Eduardo (1998) 'Competition and equalization of inter-industry profit rates: the evidence for the Brazilian economy, 1973–85', *International Working Group on Value Theory mini-conference at the Eastern Economic Association*, New York, N.Y.

- Mantegna, Rosario N. and Stanley, H. Eugene (2000) *An introduction to econophysics: correlations and complexity in finance*, Cambridge University Press, Cambridge.
- Markose, Sheri M. (2005) 'Computability and evolutionary complexity: markets as complex adaptive systems', *The Economic Journal*, **115**:June, pp. F159-F192.
- Marsaglia, G. (1974) 'Extension and application of Lukacs' characterization of the gamma distribution', *Proc. Symposium Statistics & Related Topics*, Carleton University, Ottawa.
- Marsaglia, G. (1989) 'The $X + Y$, X/Y characterization of the gamma distribution', in *Contributions to probability and statistics*, (Eds, Bleier, L.J., Perlman, M.D., Press, S.J. and Sampson, A.R.) Springer-Verlag, New York, pp. 91-98.
- Martin, Stephen (1984) 'The misuse of accounting rates of return: comment', *The American Economic Review*, **74**:3, p. 501-506.
- Marx, Karl (1952) *Wage labour and capital*, Progress Publishers, Moscow.
- Marx, Karl (1969) *Theories of surplus value*, Vol. 2, Lawrence and Wishart, London.
- Marx, Karl (1975a) 'Excerpts from James Mill's *Elements of political economy*', in *Early Writings*, (Ed, Colletti, Lucio) Penguin, Harmondsworth, pp. 259-278.
- Marx, Karl (1975b) 'Difference between the Democritean and Epicurean philosophy of nature', in *Marx-Engels collected works*, Vol. 1 Lawrence and Wishart, London.
- Marx, Karl (1976) *Capital*, Vol. I, Penguin Books, Harmondsworth.
- Marx, Karl (1978) *Capital*, Vol. II, Penguin Books, Harmondsworth.
- Marx, Karl (1981) *Capital*, Vol. III, Penguin Books, Harmondsworth.
- Marx, Karl (n.d.) *The poverty of philosophy*, Lawrence and Wishart, London.
- McAlister, D. (1879) 'The law of the geometric mean', *Proceedings of the Royal Society*, **29**, p. 367.
- McCullagh, P. and Nelder, J.A. (1989) *Generalized linear models*, Chapman and Hall, London.
- McDonald, J.B. (1984) 'Some generalized functions for the size distribution of income', *Econometrica*, **52**, pp. 647-663.
- McLellan, David (1980) *Marx before marxism*, Macmillan, London.
- Mitzenmacher, Michael (2003) 'A brief history of generative models for power law and lognormal distributions', *Internet Mathematics*, **1**:2, pp. 226-251.
- Moseley, Fred (1991) *The falling rate of profit in the postwar United States economy*, Macmillan.
- Moseley, Fred (1992) 'Unproductive labour and the rate of profit in the postwar US economy', in *International perspectives on profitability and accumulation*, (Eds, Moseley, Fred and Wolff), pp. 66-90.
- Mosselmans, Bert (2005) 'Adolphe Quetelet, the average man and the development of economic methodology', Antwerp, pp. 25.
- Parker, R.H. and Harcourt, Geoff C. (1969a) 'Introduction', in *Readings in the concept and measurement of income*, Cambridge University Press, Cambridge, p. 402.
- Parker, R.H. and Harcourt, Geoff C. (Eds.) (1969b) *Readings in the concept and measurement of income*, Cambridge University Press, Cambridge.
- Perroux, Francois (1926 [1996]) *Le probleme du profit*, Presses Universitaires de Grenoble.
- Puty, Cláudio Castelo Branco (n.d.) 'A simulation model of the general rate of profit with capital accumulation', Belém, Brasil.
- Quetelet, Lambert Adolphe Jacques (1842 [1969]) *A treatise on man and the development of his faculties*, Scholars' Facsimiles and Reprints, Gainesville, Florida.
- Ravenscraft, David J. (1983) 'Structure-profit relationships at the line of business and industry level', *The Review of Economics and Statistics*, **65**, pp. 22-31.

- Ricardo, David (1951) *On the principles of political economy and taxation*, Vol. I, Cambridge University Press, Cambridge.
- Rider, P.R. (1933) 'Criteria for rejection of observations', *Washington University Studies – New Series, Science and Technology*, 8, p. 3–23.
- Russo, Alberto, Delli Gatti, Domenico and Gallegati, Mauro (2006) 'Technological innovation, financial fragility and complex dynamics', in *Economic growth and distribution: on the nature and causes of the wealth of nations*, (Ed, Salvadori, Neri) Edward Elgar, Cheltenham.
- Salamon, G.L. (1982) 'Cash recovery rates and measures of firm profitability', *Accounting Review*, 57:April, pp. 292–302.
- Salamon, G.L. (1985) 'Accounting rates of return', *American Economic Review*, 75:June, pp. 495–504.
- Salem, A.B.Z. and Mount, T.D. (1974) 'A convenient descriptive model of income distribution: the gamma density', *Econometrica*, 42:6, pp. 1115–1128.
- Serrano Cinca, C., Mar Molinero, C. and Gillizo Larraz, J.L. (2001) 'Country and size effects in financial ratios: a European perspective', University of Southampton, Southampton.
- Shaikh, Anwar (1978) 'National income accounts and marxian categories', New School for Social Research, New York.
- Shapiro, S.S. and Chen, L. (2001) 'Composite tests for the gamma distribution', *Journal of Quality Technology*, 33:1, pp. 47–59.
- Sheppard, Eric and Barnes, Trevor J. (1986) 'Instabilities in the geography of capitalist production: collective vs. individual profit maximization', *Annals of the Association of American Geographers*, 76, pp. 493–507.
- Sheppard, Eric and Barnes, Trevor J. (1990) *The capitalist space economy: geographical analysis after Ricardo, Marx and Sraffa*, Unwin Hyman, London.
- Sherman, Howard J. (1968) *Profits in the United States*, Cornell University Press, Ithaca.
- Silverman, Bernard W. (1986) *Density estimation for statistics and data analysis*, Chapman and Hall, London.
- Simons, H.C. (1969) 'The definition of income', in *Readings in the concept and measurement of income*, (Eds, Parker, R.H. and Harcourt, Geoffrey C.) Cambridge University Press, Cambridge, pp. 63–73.
- Singh, A. and Whittington, G. (1968) *Growth, profitability and valuation*, Cambridge University Press, Cambridge.
- Stacy, E.W. and Mihram, G.A. (1965) 'Parameter estimation for a generalized gamma distribution', *Technometrics*, 7:3, pp. 349–358.
- Stanley, M.H.R., Buldyrev, S.V., Havlin, S., Mantegna, R.N., Salinger, M.A. and Stanley, H.E. (1995) 'Zipf plots and the size distributions of firm', *Economics Letters*, 49, pp. 453–457.
- Steedman, Ian (1977) *Marx after Sraffa*, NLB.
- Steedman, Ian (1983) '[Review of *Laws of Chaos*]', *The Economic Journal*:December, pp. 1015–1016.
- Steibeling, G.C. (1890) *Das Werthgesetz und die Profitrate*, John Heinrich, New York NY.
- Stigler, George J. (1963) *Capital and rates of return in manufacturing industries*, Princeton University Press, Princeton.
- Stigler, S.M. (1973) 'Simon Newcomb, Percy Daniell, and the history of robust estimation, 1885–1920', *Journal of the American Statistical Association*, 68, p. 872–879.
- Stigler, Stephen M. (1986) *The history of statistics: the measurement of uncertainty before 1900*, The Belknap Press of Harvard University Press, Cambridge, Mass.

- Stone, E.J. (1873) 'On the rejection of discordant observations', *Monthly Notices of the Royal Astronomical Society*, **34**, p. 9–15.
- Sutton, J. (1997) 'Gibrat's legacy', *Journal of Economic Literature*, **35**, pp. 40–59.
- Taguchi, Tokio (1980) 'On an interpretation and an estimation of shape parameters of the generalized gamma distribution', *Metron*, **38**, pp. 27–40.
- Tsaliki, Persefoni (1998) 'Theories of competition and concept of regulating capital: evidence from Greek manufacturing', *IWGVIT*, New York, N.Y.
- van der Borghst (1898) 'Aktiengesellschaft', in *Dictionnaire d'économie politique*, p. 191.
- Verso - NLB (Ed.) (1981) *The value controversy*, Verso.
- Webber, M.J. (1986) 'Survey 9. The theory of prices, profits and values', *Society and Space*, **4**, pp. 109–116.
- Webber, Michael (1987) 'Rates of profit and interregional flows of capital', *Annals of the Association of American Geographers*, **77**:1.
- Webber, Michael J. and Rigby, D.L. (1986) 'The rate of profit in Canadian manufacturing industry, 1950–1981', *Review of Radical Political Economics*, **18**:1 & 2, pp. 33–55.
- Webber, Michael J. and Rigby, David .L. (1996) *The Golden Age illusion: rethinking postwar capitalism*.
- Weisskopf, Thomas E. (1979) 'Marxian crisis theory and the rate of profit in the postwar US economy', *Cambridge Journal of Economics*, **3**:4, pp. 341–378.
- Weisskopf, Thomas E. (1981) 'Wages, salaries, and the profit share: a rejoinder', *Cambridge Journal of Economics*:5, pp. 175–182.
- Wolff, Edward N. (1979) 'The rate of surplus value, the organic composition, and the general rate of profit in the U.S. economy, 1947–67', *The American Economic Review*, **69**:3, pp. 329–341.
- Wolff, Edward N. (1986) 'The productivity slowdown and the fall in the U.S. rate of profit, 1947–76', *Review of Radical Political Economy*, **18**:1 & 2, pp. 87–109.
- Wolff, Edward N. and Dollar, David (1992) 'Profit and wage convergence and capital accumulation among industrialised countries, 1963–83', in *Profits, deficits and instability*, (Ed, Papadimitriou, Dimitri B.) Macmillan, London, pp. 103–138.
- Wright, Ian (2005) 'The social architecture of capitalism', *Physica A: Statistical and Theoretical Physics*, **346**:3–4, p. 589–620.
- Zachariah, Dave (2005) 'Labor value and equalization of profit rates in Sweden', Stockholm, Sweden.
- Zipf, G. (1932) *Selective studies and the principle of relative frequency in language*, Harvard University Press, Cambridge MA.